Annual Issue 2013

The Journal on Advanced Studies in Theoretical and Experimental Physics, including Related Themes from Mathematics

PROGRESS IN PHYSICS

"All scientists shall have the right to present their scientific research results, in whole or in part, at relevant scientific conferences, and to publish the same in printed scientific journals, electronic archives, and any other media."
— Declaration of Academic Freedom, Article 8

ISSN 1555-5534

The Journal on Advanced Studies in Theoretical and Experimental Physics, including Related Themes from Mathematics

PROGRESS IN PHYSICS

A quarterly issue scientific journal, registered with the Library of Congress (DC, USA). This journal is peer reviewed and included in the abstracting and indexing coverage of: Mathematical Reviews and MathSciNet (AMS, USA), DOAJ of Lund University (Sweden), Zentralblatt MATH (Germany), Scientific Commons of the University of St. Gallen (Switzerland), Open-J-Gate (India), Referativnyi Zhurnal VINITI (Russia), etc.

Electronic version of this journal:
http://www.ptep-online.com

ISSN: 1555-5534 (print)
ISSN: 1555-5615 (online)

Standard Address Number: 297-5092
Printed in the United States of America

JANUARY 2013 **VOLUME 1**

CONTENTS

Information for Authors and Subscribers

Progress in Physics has been created for publications on advanced studies in theoretical and experimental physics, including related themes from mathematics and astronomy. All submitted papers should be professional, in good English, containing a brief review of a problem and obtained results.

All submissions should be designed in LaTeX format using *Progress in Physics* template. This template can be downloaded from *Progress in Physics* home page http://www.ptep-online.com. Abstract and the necessary information about author(s) should be included into the papers. To submit a paper, mail the file(s) to the Editor-in-Chief.

All submitted papers should be as brief as possible. We accept brief papers, no larger than 8 typeset journal pages. Short articles are preferable. Large papers can be considered in exceptional cases to the section *Special Reports* intended for such publications in the journal. Letters related to the publications in the journal or to the events among the science community can be applied to the section *Letters to Progress in Physics*.

All that has been accepted for the online issue of *Progress in Physics* is printed in the paper version of the journal. To order printed issues, contact the Editors.

This journal is non-commercial, academic edition. It is printed from private donations. (Look for the current author fee in the online version of the journal.)

Some Expressions for Gravity without the Big G and their Possible Wave-Theoretical-Explanation

Hasmukh K. Tank
Indian Space Research Organization, 22/693, Krishna Dham-2, Vejalpur, Ahmedabad-380015, India
E-mail: tank.hasmukh@rediffmail.com, hasmukh.tank1@gmail.com

This letter presents some new expressions for gravity without the big G and proposes their possible wave-theoretical-explanation. This attempt leads to some insight that: (i) We need the proportionality-constant G because we measure masses and distances in our arbitrarily-chosen units of kg and meters; but if we measure "mass" as a fraction of "total-mass of the universe" M_0 and measure distances as a fraction of "radius-of-the-universe" R_0 then there is no need for the proportionality-constant G. However, large uncertainties in the M_0 and R_0 limit the general application of this relation presently. (ii) The strength of gravity would be different if the total-mass of the universe were different. Then this possibility is supported with the help of wave-theory. (iii) This understanding of G leads to an insight that Plancks-length, Planck-mass and Planck's unit of time are geometric-mean-values of astrophysical quantities like: total-mass of the universe and the smallest-possible-mass hH_0/c^2. (iv) There appears a law followed by various systems-of-matter, like: the electron, the proton, the nucleus-of-atom, the globular-clusters, the spiral-galaxies, the galactic-clusters and the whole universe; that their ratio $\text{Mass}/\text{Radius}^2$ remains constant. This law seems to be more fundamental than the fundamental-forces because it is obeyed irrespective of the case, whether the system is bound by strong-force, electric-force, or gravitational-force.

1 Introduction

Sir Isaac Newton presented the quantitative description of gravitational attraction between two massive bodies, that the force of attraction is directly proportional to the product of two masses, and inversely proportional to the square of centre-to-centre distance between them; and the value of proportionality-constant G was found to remain the same even in the case of planets. But there has been no explanation for why the value of G is this much. Einstein also made extensive use of G by treating it as a fundamental-physical-constant. Based on my previous works, [1-5] and the works of researchers cited in these papers, this paper presents some alternative expressions for gravity, without the big G, and proposes a wave-theoretical-explanation for gravity.

2 New expressions of gravity without the big G

(i) R.K. Adair, in his book "Concepts in Physics" [6] has given a derivation, that the sum of "gravitational-potential-energy" and "energy-of-mass" of the whole universe is, strikingly, zero! i.e.

$$M_0c^2 - \frac{GM_0M_0}{R_0} = 0$$

where M_0 and R_0 are total-mass and radius of the universe respectively, and G is Newton's gravitational constant; i.e.

$$\frac{GM_0^2}{R_0} = M_0c^2$$

i.e.

$$G = \frac{R_0c^2}{M_0}.$$

So, by substituting R_0c^2/M_0 for G in Newton's formula, the gravitational potential energy U_g stored in a system of masses M and m separated by a distance r can be expressed as:

$$U_g = \frac{\frac{M}{M_0}mc^2}{r/R_0}. \quad (1)$$

Newton's law when expressed as shown in the expression-1, shows that: if we measure masses as a fraction of total-mass of the universe M_0 and measure distances as a fraction of radius of the universe R_0 then we do not need the big G.

However, large uncertainties in the M_0 and R_0 limit the general application of this relation presently.

A brief discussion will be in order, how the "total-mass-of-the-universe" and "radius-of-the-universe" are derived; and what would be the uncertainties of these?

Total-mass-of-the-universe:

E.P. Hubble's experimental-observations of the "cosmological-red-shift", when interpreted in terms of "recession-of-galaxies", gives a linear relation:

$$v = H_0D$$

where: v is the "velocity-of-recession" of a galaxy, H_0 is Hubble's constant and D the luminosity-distance of a galaxy. From this relation we can get an estimate of "sum-total-of-kinetic-energy-of-the-universe" K_u. This recession-of-galaxies, also known as: "expansion-of-the-universe", can stop if

and when "kinetic-energy-of-the-universe" K_u becomes equal to "gravitational-potential-energy-of-the-universe" U_u. By equating $K_u = U_u$, cosmologists have derived the value of "total-mass-of-the-universe" M_0.

It has been estimated [7] that the universe would have collapsed to hot-death much sooner than the present-age of the universe if total-mass of the universe were more than M_0; and it would have cooled down to cold-death much earlier than the present-age of the universe if its total-mass were less than M_0. The present-age, of 14 billion years, imply that the total-mass of the universe is indeed M_0. $M_0 = 10^{82}$ pion-masses.

It is surprising [8] that cosmologists are so far able to experimentally detect only the baryonic-matter, which is hardly 4% of the total-mass M_0! At least 70% of the total-mass M_0 is believed to be in the form of "dark-energy", and remaining 26% in the form of "dark-matter". "Dark-matter" is needed to explain the "flattening-of-galaxies-rotation-curves". That is, the estimates of total-mass of the universe depend on 26% share from "dark-matter", and 70% share from "dark-energy" which are yet to be detected.

Radius-of-the-universe:

The distance at which a galaxy can attain the velocity-of-light, that is, when Hubble's expression becomes: $H_0R_0 = c$, where c is the speed-of-light, this distance R_0 is called: "the-radius-of-the-universe". Even if universe-tip may be moving with speed higher than light-speed, the "visible" horizon will be limited by the equation $c = H_0R_0$ [8]. So, the value of radius of the universe is taken as 10^{26} meters, i.e. $= 10^{40}$ classical-radius of the electron. Here H_0 is Hubble's constant.

As far as accuracy of the values of M_0 and R_0 are concerned, there must be large amount of uncertainties. We can not expect to improve current value of G form them. Our expression of gravity without G can only help us to gain an insight, that the strength of gravitational-force seems to depend on total-mass and radius of the universe. Similarly, we can gain some insight in to Planck's natural units, and Milgrom's new constant of nature a_0, termed as the "critical-acceleration" of Modified Newtonian Dynamics (MOND).

Now, let us move to some more expressions without the big G.

(ii) Milgrom's expression for the constant velocity v of the stars at the out-skirts of a spiral-galaxy of mass M is conventionally expressed as [7]:

$$v = (GMa_0)^{1/4}\,. \tag{2}$$

Since: $G = R_0c^2/M_0$, and $a_0 = c^2/R_0$, as discussed in [9], the expression-2 can be re-expressed without G as:

$$v = \left[\frac{R_0c^2}{M_0}M\frac{c^2}{R_0}\right]^{1/4}$$

i.e.

$$v = [M/M_0]^{1/4}c. \tag{3}$$

In the expression-3, c is the speed of light in vacuum, and M_0 and R_0 are total-mass and radius of the universe respectively.

(iii) We can express the radii of the globular-clusters, the spiral-galaxies and the galactic-clusters as:

$$R_{globu} = \left[\frac{M_{globu}}{M_0}\right]^{1/2}R_0 = [r_{Gglo}R_0]^{1/2}, \tag{4}$$

where r_{Gglo} is gravitational-radius of the globular-cluster.

$$R_{galaxy} = \left[\frac{M_{galaxy}}{M_0}\right]^{1/2}R_0 = [r_{Ggal}R_0]^{1/2}, \tag{5}$$

where r_{Ggal} is gravitational-radius of the galaxy.

$$R_{gal-clust} = \left[\frac{M_{gal-clust}}{M_0}\right]^{1/2}R_0 = [r_{Ggal-clust}R_0]^{1/2}, \tag{6}$$

where $r_{Ggal-clust}$ is gravitational-radius of the galactic-cluster.

Even the classical-radius of the electron $r_e = e^2/m_ec^2$ can also be expressed as:

$$r_e = \left[\frac{m_e}{M_0}\right]^{1/2}R_0 = [r_{G-e}R_0]^{1/2}, \tag{7}$$

where r_{G-e} is gravitational-radius of the electron.

Radius of the pi-meson $r_{pi} = Ng^2/m_{pi}c^2$ can also be expressed as:

$$r_{pi} = \left[\frac{m_{pi}}{M_0}\right]^{1/2}R_0 = [r_{G-pi}R_0]^{1/2}, \tag{8}$$

where r_{G-pi} is gravitational-radius of the pi-meson.

And the radius of nucleus of an atom r_n can also be expressed as:

$$r_n = \left[\frac{m_n}{M_0}\right]^{1/2}R_0 = [r_{G-n}R_0]^{1/2}, \tag{9}$$

where r_{G-n} is gravitational-radius of the nucleus-of-atom.

The expressions (4) to (9) can be jointly expressed as [8]:

$$\frac{M_0}{R_0^2} = \frac{m_P}{r_P^2} = \frac{me}{r_e^2} = \frac{m_n}{r_n^2} = \frac{M_{gc}}{R_{gc}^2} = \frac{M_{gal}}{R_{gal}^2} = \frac{M_{cg}}{R_{cg}^2} = \frac{H_0c}{G}.$$

We shall consider a possible "wave-theoretical-explanation" for the expressions (4) to (9) in the section-4.

Since the classical-radius of the electron $r_e = e^2/m_ec^2$, radius of the pi-meson $r_{pi} = Ng^2/m_{pi}c^2$ and the radius of nucleus of an atom r_n can also be expressed in the similar manner by inserting the masses of the electron, the pi-meson and the nucleus in the right-hand-sides of the above expressions, though they are bound by electric-force, strong-force and the nuclear-force respectively, it suggests a possibility that the currently-believed fundamental-forces may not be truly fundamental; rather, the law followed by them, as expressed in

the expressions (4) to (9), may be more fundamental than the "fundamental-forces"; and the strengths of forces may be getting decided by these expressions (4) to (9). It also suggests a possibility that when a "black-hole" has some "mass" then it has to have a "radius".

(iv) We can express the cosmological red-shift z_c smaller than unity as:

$$z_c = \frac{D}{R_0}. \tag{10}$$

And we can express the accelerated-expansion of the universe, the deceleration of the cosmologically red-shifted photon, the deceleration of the Pioneer-10, 11, Galileo and Ulysses space-probes and the "critical-acceleration" of MOND as [9]:

$$a_0 = \frac{c^2}{R_0}. \tag{11}$$

3 Some insight into Planck's units

From the law of equality of gravitational-potential-energy and energy-of-mass of the universe we found that: $G = R_0c^2/M_0$. Now let us make use of this expression to get some insight into Planck's units of length, mass and time:
Planck's-length $L^* = [hG/c^3]^{1/2}$.

Substituting R_0c^2/M_0 for G in the above expression,
Planck's-length $L^* = [hR_0c^2/M_0c^3]^{1/2}$;
i.e. Planck's-length

$$L^* = \left[\frac{h}{M_0c}R_0\right]^{1/2}; \tag{12}$$

i.e. Planck's-length L^* is a geometric-mean of: Compton-wavelength and Gravitational radius of total-mass of the universe, because $R_0 = GM_0/c^2$.
Planck-mass $M^* = [hc/G]^{1/2}$;
i.e. Planck-mass $M^* = [(h/R_0c)(M_0)]^{1/2}$; i.e. Planck-mass

$$M* = \left[\frac{hH_0}{c^2}M_0\right]^{1/2}. \tag{13}$$

That is Planck's unit of mass is a geometric-mean of: total-mass of the universe and smallest-possible-mass, corresponding to Hubble's constant (hH_0/c^2). Similarly, Planck's unit of time T^* is a geometric-mean of: age-of-the-universe T_0 and the period (h/M_0c^2): i.e.

$$T^* = \left[T_0\frac{h}{M_0c^2}\right]^{1/2}. \tag{14}$$

4 Possible wave-theoretical explanation for gravity

Let us assume that there are some most-fundamental-particles, and a long-range fundamental-force. We can take the mass of the "most-fundamental-particle" as a unity, and think that all the massive objects are collections of the "most-fundamental-particles".

Now, by a "particle" we mean an entity which is localized in an extremely small space; so, a "particle" can be mathematically represented in the space-domain as an impulse-function. This impulse-function can be Fourier-transformed into the "wave-number-domain". Then assuming a constant velocity of transmission of these waves, at the velocity of light, we can represent these waves in the "frequency-domain" as a wide band of frequencies. A particle of matter has a wide band of frequency-spectrum and a definite phase-spectrum. When this wide band of waves travels in space, then a "particle" becomes manifest only at a place and time when-and-where all the spectral-components add constructively, and have a particular, definite phase-relation, otherwise the particle remains dissolved in the un-manifest-state.

Secondly, we can not expect any coherence between the spectral-components of one and the other "particle". That means, that when two or more such fundamental-particles come close to each-other, the wide bands of their waves add like the incoherent superimposition of wideband-noise.

We know that the superimposition of n number of wide-band noise-sources of unit-amplitude is square-root-of n; like the vector-sum of n mutually orthogonal unit-vectors. That is:

$$N(t) = \left[(N_1(t))^2 + (N_2(t))^2 + (N_3(t))^2 \cdots + (N_n(t))^2\right]^{1/2}.$$

Now, if the strength of "coupling-constant" of a fundamental-force is, say, e^2, which is the strength of electric-force of the proton, then the strength of "coupling-constant" of a new "fundamental-force", which is actually due to "incoherent-superimposition", within the system of n fundamental-particles will be: $(n^{1/2}e^2)/n$. Since the total-mass of the universe M_0 is 10^{80} proton-masses, the strength of gravitational-force between the two protons is expected to be:

$$GM_0m_p = (\text{Total}-\text{number of protons in the universe})^{1/2}\ e^2$$

i.e.

$$Gm_p^2 = \frac{\sqrt{10^{80}}e^2}{10^{80}}$$

i.e.

$$Gm_p^2 = 10^{-40}e^2. \tag{15}$$

[Note: This is just an order-of-magnitude-estimate.]

Now, if the force within a system is stronger than gravity by a multiplication-factor, say, k-times, then the density of matter within that system is also logically expected to be k-times higher. That is, in our example of proton and the universe:

$$\frac{e^2}{Gm_p^2} = \left[\frac{M_0}{m_p}\right]^{1/2} = \frac{\frac{m_p}{\frac{4}{3}\pi r_p^3}}{\frac{M_0}{\frac{4}{3}\pi R_0^3}}$$

i.e.

$$\frac{e^2}{Gm_p^2} = \left[\frac{M_0}{m_p}\right]^{1/2} = \frac{m_pR_0^3}{M_0r_p^3}$$

i.e.

$$\left[\frac{M_0}{m_p}\right]^{3/2} = \frac{R_0^3}{r_p^3}$$

i.e.

$$\left[\frac{M_0}{m_p}\right]^{1/2} = \frac{R_0}{r_p} = \frac{e^2}{Gm_p^2} \tag{16}$$

i.e.

$$\frac{M_0}{R_0^2} = \frac{m_p}{r_p^2}. \tag{17}$$

The expression-16 was noticed as the "large-number-coincidence" [LNC], whereas here we derived it with the help of wave-theory.

Sivaram [10] had noticed a relation between masses and radii of the electron, the proton the nucleus-of-atoms, the globular-clusters, the spiral-galaxies, the galactic-clusters and the universe as shown in the expression-18 below. The expression-18 is similar to the expression-17 derived by us using wave-theory. So our derivation based on wave-theory matches with the observations presented by Sivaram.

$$\frac{M_0}{R_0^2} = \frac{M_{gal-clust}}{R_{gal-clust}^2} = \frac{M_{gal}}{R_{gal}^2} = \frac{M_{globu}}{R_{globu}^2} = \frac{m_n}{r_n^2} = \frac{m_{pi}}{R_{pi}^2} = \frac{m_e}{r_e^2} = \frac{H_0 c}{G} \tag{18}$$

Even the mysterious-looking Weinberg-formula can be re-written, and explained, as follows: Weinberg's formula is: $m_{pi}^3 = h^2 H_0/c\, G$, which can be re-written as: $m_{pi}/(h/m_{pi}c)^2 = H_0 c/G$. Weinberg's formula has an imbalance of one order of magnitude which can be corrected by replacing Compton-wavelength of the pion by radius of the pion, i.e. $m_{pi}/R_{pi}^2 = H_0 c/G$. So the mysterious-looking Weinberg-formula is also a part of the expression-18.

5 Conclusion

Now we have an explanation for why we need the gravitational constant G. The strength of gravity seems to depend on the total-mass M_0 and radius R_0 of the universe. However, large uncertainties in the M_0 and R_0 limit the general application of this relation presently. Secondly, gravity may not be an independent "fundamental-force"; it may be arising due to "in-coherent super-imposition" of wave-amplitudes of very wide-band of waves of total number of fundamental-particles contained in the universe. The theory also explained the large-number-coincidence, and the mysterious-looking Weinberg formula. We also gained some insight into Planck's units that: Planck-length, Planck-mass and Planck's unit of time are geometric-mean-values of astrophysical quantities like: total-mass of the universe and the smallest-possible-mass hH_0/c^2.

Submitted on: September 20, 2012 / Accepted on: September 28, 2012

References

1. Tank H.K. Explanation for the recurrences of large-number 10^{40} in astrophysics, and some insight into the nature of fundamental forces. *Proceedings of Indian National Science Academy*, 1997, v. 63A, no. 6, 469–474.
2. Tank H.K. An Explanation for the Relative Strengths of 'Gravitational' and 'Electric' Forces Suggesting Equality of the 'Electrostatic-potential-energy', 'Gravitational-potential-energy' and 'Energy of Mass' of the Universe. *Science and Culture*, 2009, v. 75 (searchable from google).
3. Tank H.K. A new law emerging from the recurrences of the "critical-acceleration" of MOND, suggesting a clue to unification of fundamental forces. *Astrophysics and Space Science*, 2010, v. 330, 203–205.
4. Tank H.K. Wave-theoretical explanation for the newly-emerged-law of equality of potential-energy and energy-of-mass of reasonably independent systems of matter. *Advanced Studies in Theoretical Physics*, 2011, v. 5, 45–55.
5. Tank H.K. Some clues to understand MOND and the accelerated expansion of the universe. *Astrophysics and Space Science*, 2011, v. 336, no. 2, 341–343.
6. Adair R.K. Concepts in Physics. Academic Press, New York, (1969), p. 775.
7. Gupta R.C. and Pradhan, Anirudh. A Theoretically-explained New-variant of Modified Newtonian Dynamics (MOND). *Astrophysics and Space Science*, 2011, v. 333, 311–316.
8. Rees M. Just Six Numbers. Orion Books, London, (2000), p. 81–82.
9. Tank H.K. Genesis of the critical-acceleration of MOND and its role in formation of structures. *Progress in Physics*, 2012, v. 4, 36–38.
10. Sivaram C. Some aspects of MOND and its consequences for cosmology. *Astrophysics and Space Science*, 1994, v. 215, 185–189.

A Theoretical Description of U(5)-SU(3) Nuclear Shape Transitions in the Interacting Boson Model

A.M. Khalaf* and T.M. Awwad†

*Physics Department, Faculty of Science, Al-Azhar University, Egypt. E-mail: Ali-Khalaf43@hotmail.com
†Department of Physics, Faculty of Girls, Ain Shams University, Egypt. E-mail: tawwad12@hotmail.com

We investigated the evaluation of nuclear shape transition from spherical to axially rotational shapes using the Coherent state formalism of the first version of interacting boson model (sd IBM). The validity of such model is examined for rare-earth Nd/Sm/Gd/Dy isotopic chains by analyzing the potential energy surface (PES's). In this region, a change from spherical to well-deformed nuclei is observed when moving from the lighter to heavier isotopes.

1 Introduction

In recent years, the study of quantum phase transition (QPT) is an important topic in the research of nuclear structure. Some evidence of nuclear shape transition have been observed. For instance, several isotopes have been found to undergo shape phase evolution of first order from spherical vibrator to deformed axially symmetric rotor and phase transition of second order from spherical vibrator to deformed $\gamma-$ soft [1–3].

The Hamiltonian describing this transition is a repulsive boson pairing Hamiltonian that has the particularity of being exactly solvable allowing the study of very large systems. The study of phase shape transitions in nuclei can be best done in the interacting boson model (IBM) [4] which reproduces well the data in all transition regions [5–11].

The possible phases that can occur in the IBM have been classified in a triangular Casten diagram [12], the three phases correspond to the breaking of U(6) into its three subalgebras U(5), SU(3) and O(6) [13]. The $X(5)$ critical point symmetry [14] was developed to describe analytically the structure of nuclei at the critical point of the transition from vibrational U(5) to prolate axially symmetric SU(3) shapes. In addition the symmetry E(5) [15, 16] have been introduced to describe the nuclei at the critical point corresponding to second order transition, nuclear examples of which were used [17]. Recently, the critical point in the phase transition from axially deformed to triaxial nuclei called Y(5), has been analyzed [18]. In all these cases, critical points are defined in the context of the collective Bohr Hamiltonian [19].

Since the IBM was formulated from the beginning in terms of creation and annihilation boson operators, its geometric interpretation in terms of shape variables is usually done by introducing a boson condensate with two shape parameters β and γ. The parameter β is related to the axial deformation of the nucleus, while γ measures the deviation from axial symmetry. The equilibrium shape of the nucleus is obtained by minimizing the expectation value of the Hamiltonian in the intrinsic state.

In this paper, we discuss some aspects of the nuclear shape phase transition in even-even nuclei using the IBM with the intrinsic state formalism. The outline of the present paper is as follows: In Section 2, we construct the IBM Hamiltonian in terms of Casimir operators and using coherent state to get the potential energy surface (PES). In section 3, we check that results of the IBM with coherent state to agree for dynamical limits U(5), SU(3) and O(6) in the limit of large N. In section 4 we applied our model to the rare earth Nd/Sm/Gd/Dy isotopic chains which evolve a rapid structural changes from spherical to well-deformed nuclei when moving from lighter to the heavier isotopes.

2 Coherent State Potential Energy Surface

We start by considering a general standard two-body sd IBM Hamiltonian in the Casimir forms as:

$$\begin{aligned} H &= \epsilon C_1[U(5)] + K_1 C_2[U(5)] \\ &\quad + K_2 C_2[O(5)] + K_3 C_2[O(3)] \\ &\quad + K_4 C_2[SU(3)] + K_5 C_2[O(6)] \end{aligned} \tag{1}$$

Here $C_n[G]$is the n-rank Casimir operator of the Lie group G, with

$$C_1[U(5)] = \hat{n}_d \tag{2}$$

$$C_2[U(5)] = \hat{n}_d(\hat{n}_d + 4) \tag{3}$$

$$C_2[O(5)] = 4[\frac{1}{10}(\hat{L}\,\hat{L}) + \hat{T}_3\,\hat{T}_3] \tag{4}$$

$$C_2[O(3)] = 2(\hat{L}\,\hat{L}) \tag{5}$$

$$C_2[SU(3)] = \frac{2}{3}\left[2\left(\hat{Q}\,\hat{Q}\right) + \frac{3}{4}(\hat{L}\,\hat{L})\right] \tag{6}$$

$$C_2[O(6)] = 2\left[N(N+4) - 4(\hat{P}\,\hat{P})\right] \tag{7}$$

where $\hat{n}_d, \hat{P}, \hat{L}, \hat{Q}, \hat{T}_3$ and $\hat{T}_4$ are the boson number, pairing, angular momentum, quadrupole, octupole and hexadecapole operators defined as:

$$\hat{n}_d = (d^\dagger\,\tilde{d})^{(0)} \tag{8}$$

$$\hat{P} = \frac{1}{2}(\tilde{d}\,\tilde{d}) - \frac{1}{2}(\tilde{s}\,\tilde{s}) \tag{9}$$

$$\hat{L} = \sqrt{10}\left[d^{\dagger}\times\tilde{d}\right]^{(1)} \tag{10}$$

$$\hat{T}_3 = \left[d^{\dagger}\times\tilde{d}\right]^{(3)} \tag{11}$$

$$\hat{T}_4 = \left[d^{\dagger}\times\tilde{d}\right]^{(4)} \tag{12}$$

where $s^{\dagger}(s)$and $d^{\dagger}(\tilde{d})$ are monopole and quadrupole boson creation (annihilation) operators, respectively. The scalar product is defined as

$$\hat{T}_L\,\hat{T}_L = \sum_M (-1)^M \hat{T}_{L,M}\,\hat{T}_{L,-M} \tag{13}$$

where $\hat{T}_{L,M}$ corresponds to the M component of the operator $\hat{T}_L$. The operator $\tilde{d}_m(-1)^m d_{-m}$ and $\tilde{s} = s$ are introduced to ensure the correct tensorial character under spatial rotations.

The Connection between the IBM, PES, geometric shapes and phase transitions can be investigated by introducing a coherent, or intrinsic state which is expressed as a boson condensate [20]

$$|N,\beta,\gamma\rangle = \frac{1}{\sqrt{N!}}\left(b_c^{\dagger}\right)^N|0\rangle \tag{14}$$

with

$$b_c^{\dagger} = \frac{1}{\sqrt{1+\beta^2}}\left(s^{\dagger}+\beta\cos\gamma\; d_o^{\dagger} + \frac{1}{\sqrt{2}}\beta\sin\gamma(d_2^{\dagger}+d_{-2}^{\dagger})\right). \tag{15}$$

$|0\rangle$ is the boson vacuum and the variables β and γ determine the geometry of nuclear surface. Spherical shapes are characterized by $\beta = 0$ and deformed ones by $\beta > 0$. The angle γ allows one to distinguish between axially deformed nuclei $\gamma = 0^{\circ}$ for prolate and $\gamma = 60^{\circ}$ for oblate deformation and triaxial nuclei $0^{\circ} < \gamma < 60^{\circ}$.

The expectation values of the Casimir operators equations (2–7) in the ground state equation (14) is:

$$\langle C_1[U(5)]\rangle = \frac{N}{1+\beta^2}\beta^2 \tag{16}$$

$$\langle C_2[U(5)]\rangle = \frac{5N}{1+\beta^2}\beta^2 + \frac{N(N-1)}{(1+\beta^2)^2}\beta^4 \tag{17}$$

$$\langle C_2[O(5)]\rangle = \frac{8N}{1+\beta^2}\beta^2 \tag{18}$$

$$\langle C_2[O(3)]\rangle = \frac{12N}{1+\beta^2}\beta^2 \tag{19}$$

$$\begin{aligned}\langle C_2[SU(3)]\rangle &= \frac{20}{3}N + \frac{4}{3}\frac{N(N-1)}{(1+\beta^2)^2}\cdot\\ &\left(4\beta^2 + \frac{1}{2}\beta^4 + 2\sqrt{2}\beta^3\cos(3\gamma)\right)\end{aligned} \tag{20}$$

$$\langle C_2[O(6)]\rangle = 2N(N+4) - \frac{1}{2}\frac{N(N-1)}{(1+\beta^2)^2}\left(1-\beta^2\right)^2. \tag{21}$$

The PES associated with the IBM Hamiltonian of equation (1) is given by its expectation value in the coherent state and can be written as:

$$\begin{aligned}V(\beta,\gamma) &= a_{\circ}\frac{N}{1+\beta^2}\beta^2\\ &+\frac{N(N-1)}{(1+\beta^2)^2}\left(a_1 + a_2\beta^2 +\right.\\ &\left. a_3\beta^3\cos(3\gamma) + a_4\beta^4\right)\end{aligned} \tag{22}$$

where the coefficients a_i are linear combinations of the parameters of the Hamiltonian and terms which do not depend on β and/or γ have not been included.

3 Shape Structure of the Dynamical Symmetries

The analysis of the three dynamical symmetry limits of the IBM provides a good test of the formalism presented in the previous section.

3.1 The U(5) Symmetry

The Hamiltonian of the vibrational limit $U(5)$ can be written down by putting $k_4 = k_5 = 0$ in equation (1). This has the consequence that in H remain only the terms which conserve both the number of d-bosons and the one of the s-bosons. The Hamiltonian operator of this approximation reads:

$$\begin{aligned}H[U(5)] &= \epsilon C_1[U(5)] + K_1C_2[U(5)] +\\ &K_2C_2[O(5)] + K_3C_2[O(3)].\end{aligned} \tag{23}$$

This yields the PES

$$E(N,\beta) = \epsilon_d\frac{N}{1+\beta^2}\beta^2 + f\frac{N(N-1)}{(1+\beta^2)^2}\beta^4. \tag{24}$$

This energy functional is $\gamma-$ independent and has a minimum at $\beta = 0$, Special case for $U(5)$ limit, when

$$H = \epsilon C_1[U(5)], \tag{25}$$

$$E(N,\beta) = \epsilon\frac{N}{1+\beta^2}\beta^2. \tag{26}$$

3.2 The SU(3) Symmetry

In the parametrization equation (1), the $SU(3)$ limit corresponds to $\epsilon = K_1 = K_2 = K_5 = 0$ and the Hamiltonian reads:

$$H[SU(3)] = K_3C_2[O(3)] + K_4C_2[SU(3)]. \tag{27}$$

This yields the PES

$$\begin{aligned}&E(N,\beta,\gamma) =\\ &3(4k_3+k_4)\frac{N}{1+\beta^2}\beta^2 + \frac{4}{3}k_4\left[\frac{N}{1+\beta^2}\left(5+\frac{11}{4}\beta^2\right)\right.\\ &\left.+\frac{N(N-1)}{(1+\beta^2)^2}\left(4\beta^2 + 2\sqrt{2}\beta^3\cos(3\gamma) + \frac{1}{2}\beta^4\right)\right].\end{aligned} \tag{28}$$

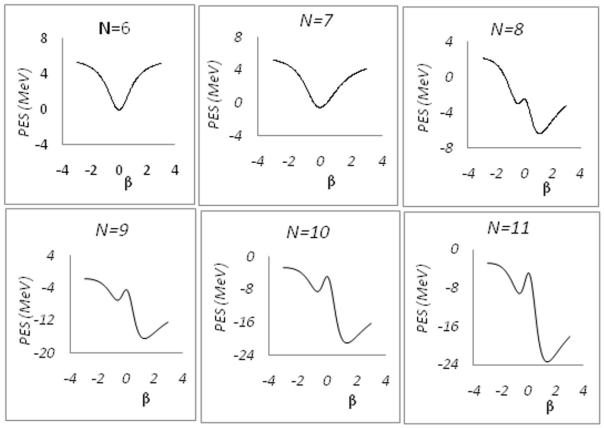

Fig. 1: Calculated PES's as a function of the deformation parameter β in U(5)-SU(3) transition for $^{144-154}$Nd (with $N_\pi = 5$ and $N_\nu = 1-6$ neutron bosons) isotopic chain. The total number of bosons N=6-11 and $\chi = -\sqrt{7}/2$).

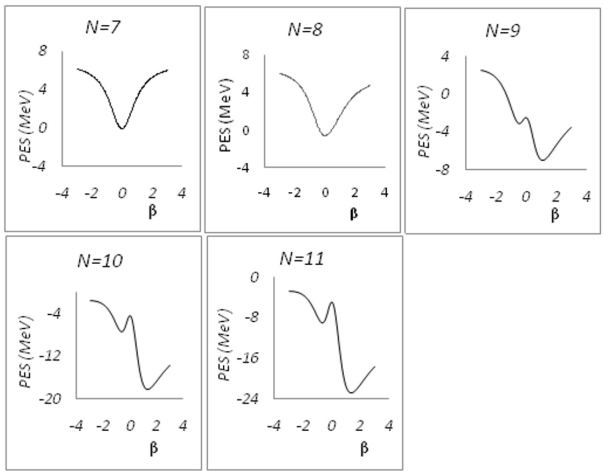

Fig. 2: Calculated PES's as a function of the deformation parameter β in U(5)-SU(3) transition for $^{146-154}$Sm (with $N_\pi = 6$ and $N_\nu = 1-5$) isotopic chain. The total number of bosons N=6-11 and $\chi = -\sqrt{7}/2$).

This energy functional has a shape minimum at $\gamma = 0$ and at a value $\beta \neq 0$.

Special case for SU(3) limit, when

$$H = a\hat{Q}\,\hat{Q} \tag{29}$$

and if we eliminate the contribution of the one-body terms of the quadrupole -quadrupole interaction, then, the PES reads

$$E(N,\beta,\gamma) = a\frac{N(N-1)}{(1+\beta^2)^2}(4\beta^2 \pm 2\sqrt{2}\beta^3\cos(3\gamma) + \frac{1}{2}\beta^4). \tag{30}$$

The equilibrium values are obtained by solving

$$\frac{\partial E}{\partial \beta} = \frac{\partial E}{\partial \gamma} = 0 \tag{31}$$

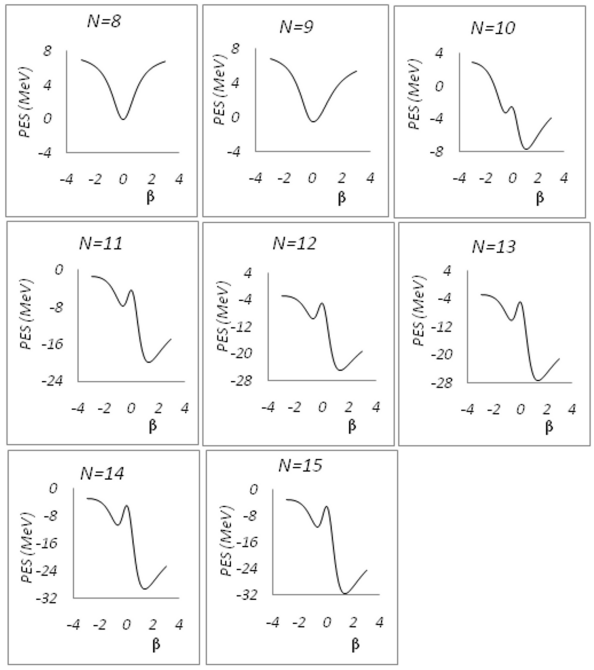

Fig. 3: Calculated PES's as a function of the deformation parameter β in U(5)-SU(3) transition for $^{148-162}$Gd (with $N_\pi = 7$ and $N_\nu = 1-8$) isotopic chain. The total number of bosons N=6-11 and $\chi = -\sqrt{7}/2$).

to give $\beta_e = \sqrt{2}$ and $\gamma = 0°$ and $\gamma = 60°$.

3.3 The O(6) Symmetry

For the O(6) limit $\epsilon = K_1 = K_2 = 0$ and the Hamiltonian takes the form

$$H[O(6)] = K_2C_2[O(5)] + K_3C_2[O(3)] + K_5C_2[O(6)]. \tag{32}$$

One then obtains the PES

$$\begin{aligned} E(N,\beta) &= 12(2K_2+K_3)\frac{N}{1+\beta^2}\beta^2 - \\ &\quad 2k_5N(N-1)\left(\frac{1-\beta^2}{1+\beta^2}\right)^2. \end{aligned} \tag{33}$$

This energy functional is $\gamma-$independent and has a minimum at a value $|\beta| \neq 0$. For large N, the minimum is at $|\beta| = 1$.

Special case for $O(6)$ limit, when

$$H = a\hat{Q}(\chi)\,\hat{Q}(\chi) \tag{34}$$

$$\chi = 0 \tag{35}$$

and if we eliminate the contribution of the one-body term of the quadrupole-quadrupole interaction, then

$$E(N,\beta) = 4aN(N-1)\left(\frac{\beta}{1+\beta^2}\right)^2 \tag{36}$$

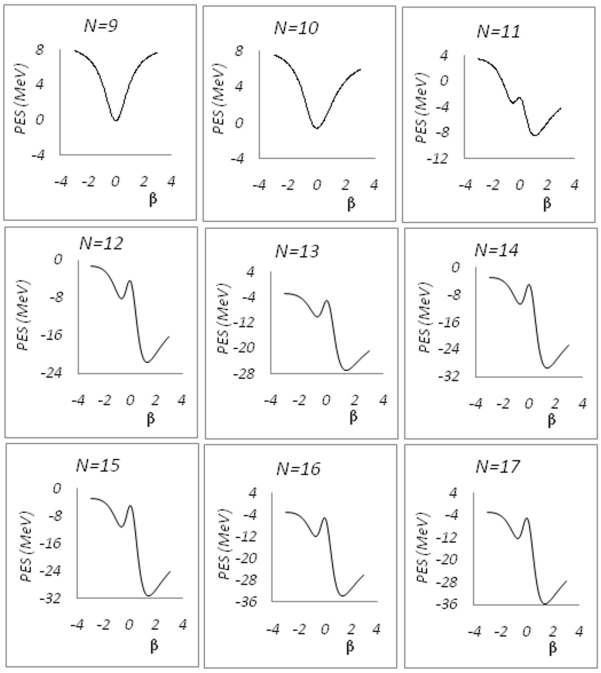

Fig. 4: Calculated PES's as a function of the deformation parameter β in U(5)-SU(3) transition for $^{150-166}$Dy (with $N_\pi = 8$ and $N_\nu = 1-9$) isptopic chain. The total number of bosons N=6-11 and $\chi = -\sqrt{7}/2$).

the equilibrium value is given by $\beta = 1$ corresponding to a γ-unstable deformed shape.

4 Application to Rare-Earth Isotope Chains

Nuclei in the region of Sm are well known examples of U(5)-SU(3) transition going from a vibrational into a rotational behavior. The validity of our model is examined for typical various even-even Nd/Sm/Gd/Dy isotopic chains with total number of bosons from N=6 to N=17.

The set of parameters of the model for each nucleus are adjusted by using a computer simulated search program in order to describe the gradual change in the structure as boson number is varied and to reproduce the properties of the selected states of positive parity excitation ($2_1^+, 4_1^+, 6_1^+, 8_1^+, 0_2^+$, $2_3^+, 4_3^+, 2_2^+, 3_1^+$ and 4_2^+) and the two neutron separation energies of all isotopes in each isotopic chain. The best fitting parameters obtained for each nucleus are given explicitly in Tables (1,2).

The PES's versus deformation parameter β for rare earth isotopic chain of nuclei evolving from spherical to axially symmetric well deformed nuclei are illustrated in figures (1-4). A first order shape phase transition with changes in number of bosons when moving from the lighter to heavier isotopes i.e. U(5)-SU(3) transitional region are observed. In our selected region we assumed a value $\chi = -\sqrt{7}/2$ because some *Gd* isotopes clearly exhibit the character of the SU(3) dynamical symmetry. Around $N = 90$ these seems to be the $X(5)$critical point symmetry. Each PES displays a relatively similar shape with only a small increase in the sharpness of the potential for increasing boson number.

5 Conclusion

In conclusion, the paper is focused on the properties of quantum phase transition between spherical U(5) and prolate deformed SU(3) in framework of the simple version of interacting boson model IBM-1 of nuclear structure.

The Hamiltonian was studied in the three different limits of the IBM and formed by laking. A systematic study of rare earth Nd/Sm/ Gd/Dy isotope chains was done using the coherent states. Nuclei located at or very close to the first order transition were the N=90 isotones ^{150}Nd, ^{152}Nd, ^{154}Nd and ^{156}Nd. They also follow the $X(5)$ pattern in ground state energies. The geometric character of the nuclei was visualizes by plotting the potential energy surface (PES's). parameters of our model were adjusted for each nucleus by using a computer simulated search program, while the parameter X in the quadrupole operator was restricted to fixed value $x = -\sqrt{7}/2$.

Submitted on: September 24, 2012 / Accepted on: September 27, 2012

References

1. Jolie J. et al. Triple Point of Nuclear Deformation. *Physical Review Letters*, 2002, v. 89, 182502–182504.
2. Arios J.M., Dukelsky J., and Garcia-Ramos J.E. Quantum Phase Transitions in the Interacting Boson Model: Integrability, Level Repulsion and Level Crossing. *Physical Review Letters*, 2003, v. 91, 162502–162504.
3. Turner P.S. and Rowe D.J. Phase Transitions and Quesidynamical Symmetry in Nuclear Collective Model, II. The Sperical vibrator to gamma-soft rotor Transition in an O(5)-invariant Bohr Model. *Nuclear Physics*, 2005, v. A756, 333–355.
4. Iachello F. and Arima A. The Interacting Boson Model. Cambridge University Press, Cambridge, England, 1987.
5. Scholten O., Iachello F. and Arima A. Interacting Boson Model of Collective Nuclear State III. The Transition from SU(5) to SU(3). *Annals of Physics*, 1978, (N.Y.) v. 115, 325–366.
6. Castonos O., Frank A. and Federman, A. The shape transition in the Sm isotopes and the structure of the IBA hamiltonian. *Physics Letters*, 1979, v. B88, 203–206.
7. Cejnar P. and Jolie J. Quantum Phase Transition Studied within the Interacting Boson Model, *Physical Review*, 2000, v. E61, 6237–6247.
8. Cejnar P., Heinze S., and Jolie J. Ground-State Shape Phase Transitions in Nuclei: Thermodynamic analogy and fitie N-effects. *Physical Review*, 2003, v. C68, 034326–034326.
9. Rowe D.J. and Thiamova G. The many relationships between the IBM and the Bohr Model. *Nuclear Physics*, 2005, v. A760, 59–81.
10. Cejnar P., Heinze S. and Dobes J. Thermodynamic analogy for quantum phase transitions at zero temperature. *Physical Review*, 2005, v. C71, 011304R–011309R.
11. Heinze S. et al. Evolution of spectral properties along the O(6)-U(5) transition in the interacting boson model. I. Level dynamics. *Physical Review*, 2006, v. C73, 014306–014316.
12. Casten R.F. Nuclear Structure from a Simple Prespective. Oxford University, Oxford, 1990.

Table 1: Values of the parameter a_o and the total number of boson for the Nd/Sm/Gd/Dy isotopic chain.

No. of Neutrons	$_{60}$Nd	$_{62}$Sm	$_{64}$Gd	$_{66}$Dy
84	1161.91775(6)	1112.0059(7)	1130.70265(8)	1174.6685(9)
86	1082.31775(7)	1078.2059(8)	1160.10265(9)	1223.4685(10)
88	1121.01775(8)	974.0059(9)	1060.40265(10)	1178.2685(11)
90	1078.51735(9)	895.5059(10)	951.80265(11)	1119.7685(12)
92	1011.71775(10)	843.3059(11)	872.90265(12)	1076.8685(13)
94	1071.51775(11)	877.2059(12)	825.80265(13)	1043.5685(14)
96	−−	996.9059(13)	813.40265(14)	1029.2685(15)
98	−−	1136.7059(14)	827.90265(15)	1025.0685(16)
100	−−	−−	−−	1058.9685(17)

Table 2: Values of the parameters a_1, a_2, a_3, and a_4 describing the IBM Hamiltonian for Nd/Sm/Gd/Dy isotopic chains.

Isotopic chains	a_1	a_2	a_3	a_4
$^{144-154}$Nd	20.93825	−110.4805	−48.51035	−84.10182
$^{146-160}$Sm	13.30225	−85.3005	−41.50433	−61.52960
$^{148-162}$Gd	11.30175	−75.1195	−37.13441	−75.61475
$^{150-166}$Dy	9.66275	−73.8775	−38.57408	−81.01053

13. Arima A. and Iachello F. Interacting Boson Model of Collective States I. The Vibrational Limit. *Annals of Physics*, 1976, (N.Y.) v. 99, 253–317; Arima A. and Iachello F. Interacting Boson Model of Collective States II. The Rotational Limit. *Annals of Physics*, 1978, (N.Y.) v. 111, 201–238.
14. Iachello F. Analytic Prescription of Critical Point Nuclei in a Spherical Axially Deformed Shape Phase Transtion. *Physical Review Letters*, 2001, v. 87, 052502–052506.
15. Iachello F. Dynamic Symmetrcies at the Critical Point. *Physical Review Letters*, 2000 v. 85, 3580–3583.
16. Leviatan A. and Ginocchio J.N. Critical Point Symmetry in a Finite System. *Physical Review Letters*, 2005, v. 90, 212501–212505.
17. Casten R.F. and Zamfir N.V. Evidence for a Possible E(5) Symmetry in ^{134}Ba. *Physical Review Letters*, 2000 v. 85, 3584–3586, Casten R.F. and Zamfir N.V. Empirical Realization of a Critical Point Description in Atomic Nuclei. *Physical Reviw Letters*, 2001 v. 87, 052503–052507.
18. Iachello F. Phase Transition an Angle Variables. *Physical Review Letters*, 2003, v. 91, 132502–132507.
19. Bohr A. and Mottelson, Nuclear Structure, Benjamin, New York, 1975, Vol. II.
20. Dieperink A.E.L., Scholten O. and Iachello F. Classical Limit of the Interacting Boson Model. *Physical Reviw Letters*, 1980, v. 44, 1747–1750.
21. Dieperink A.E.L and Scholten O. On Shapes and Shase Transition in the Interacting Boson Model. *Nuclear Physics*, 1980, v. A346, 125–138.

Fine Structure Constant as a Mirror of Sphere Geometry

Janez Špringer
Cankarjeva cesta 2, 9250 Gornja Radgona, Slovenia, EU. E-mail: janez_springer@t-2.net

A path is defined as the vector's sum of the translation and rotation component of the length unit belonging to the mass entity in motion on the sphere. The fine structure constant is an irrational number being a mirror of the path complexity as well as the sphere curvature where the path is made. The inverse value in the Euclidean plane yields $\alpha^{-1} = \sqrt{\pi^2 + 137^2}$. The inverse fine structure constant on the elliptic sphere is smaller and on the hyperbolic sphere is greater. The electron in the Hydrogen atom should move on the elliptic sphere of the radius of 3679 Compton wavelengths of the electron according to the CODATA 2012 recommended empirical value $\alpha^{-1} = 137.035999074$. Such a small sphere radius implies the heterogeneous curvature of the present universe.

1 Theoretical background

In motion is an entity having some mass. Respecting Compton the length unit is attributed to that mass:

$$\lambda = \frac{h}{mc} = 1\,. \tag{1}$$

The infinite mass and zero length unit are objectively unreachable. Nevertheless both can be theoretically approached arbitrarily close by the sufficiently great finite mass.

A curved motion obeys the path complexity: it has the translation and rotation component. Describing the curved path the length unit becomes not only the translation unit but the rotation unit, too. By the circumference of a circle concluded path s, for instance, only apparently equals the translation n, actually it is greater for the average rotation π made around the start point of the length unit:

$$\pi = \frac{0 + 2\pi \times 1}{2}\,. \tag{2}$$

The actual path is the vectorial sum of both components: the rotation π as well as translation n:

$$\overrightarrow{s} = \overrightarrow{\pi} + \overrightarrow{n}\,. \tag{3}$$

The total rotation of the length unit π equals the total Berry phase at spin $\frac{1}{2}$ [1].

1.1 Path in the Euclidean plane

By the circumference of a circle concluded path s in the Euclidean plane is calculated with the help of Pythagoras' theorem:

$$s^2 = \pi^2 + n^2. \tag{4}$$

1.2 Path on the elliptic sphere

By the circumference of a circle concluded path s on the elliptic sphere is calculated with the help of the spherical law of cosines.

On the elliptic sphere of radius R holds:

$$\cos\frac{s}{R} = \cos\frac{\pi}{R}\,\cos\frac{n}{R}\,, \tag{5}$$

$$\cos x = \sqrt{1 - \sin^2 x}\,, \tag{6}$$

$$\frac{1}{R^2} = \frac{1}{k_1^2\pi^2} + \frac{1}{k_2^2 n^2} - \frac{k_3^2 s^2}{k_1^2\pi^2 \times k_2^2 n^2} = \frac{k_1^2\pi^2 + k_2^2 n^2 - k_3^2 s^2}{k_1^2\pi^2 \times k_2^2 n^2}\,. \tag{7}$$

The coefficients are expressed as

$$k_1 = \frac{\sin\frac{\pi}{R}}{\frac{\pi}{R}}\,, \quad k_2 = \frac{\sin\frac{n}{R}}{\frac{n}{R}} \quad \text{and} \quad k_3 = \frac{\sin\frac{s}{R}}{\frac{s}{R}}\,. \tag{8}$$

They are arranged by size

$$1 \geqslant k_1 \geqslant k_2 \geqslant k_3\,. \tag{9}$$

In the case of R^2 being a positive number Pythagoras' theorem holds only exceptionally. The next condition has to be satisfied:

$$k_1^2\pi^2 + k_2^2 n^2 \geqslant k_3^2 s^2 \quad \text{or} \quad \frac{k_1^2}{k_3^2}\,\pi^2 + \frac{k_2^2}{k_3^2}\,n^2 \geqslant s^2. \tag{10}$$

The ratios of coefficients $\frac{k_1^2}{k_3^2}$ and $\frac{k_2^2}{k_3^2}$ are according to (non) equation (9) greater than 1 or at least equal 1, therefore we write:

$$\frac{k_1^2}{k_3^2}\,\pi^2 + \frac{k_2^2}{k_3^2}\,n^2 \geqslant \pi^2 + n^2 \geqslant s^2. \tag{11}$$

At the finite elliptic sphere radius R Pythagoras' theorem fails, because at non-equal coefficients (9) the square area upon hypotenuse is smaller than the sum of square areas upon catheters:

$$s^2 < \pi^2 + n^2. \tag{12}$$

At $R = \infty$ and equal coefficients (9) the elliptic sphere transforms into the Euclidean plane and Pythagoras' theorem begins to rule again (4).

1.2.1 Approximation for cos *x*

Hardy's approximation [2] is close to the function $\cos\frac{t}{r}$:

$$H\left(\frac{2t}{\pi R}\right) = \cos\frac{t}{R} \approx 1 - \frac{\left(\frac{2t}{\pi R}\right)^2}{\frac{2t}{\pi R} + \left(1 - \frac{2t}{\pi R}\right)\sqrt{\frac{2 - \frac{2t}{\pi R}}{3}}}. \tag{13}$$

At very large R Hardy's approximation can be simplified:

$$H\left(\frac{2t}{\pi R}\right) = \cos\frac{t}{R} \approx 1 - \left(\frac{2t}{\pi R}\right)^2. \tag{14}$$

The spherical law of cosines (5) with the help of the simplified Hardy approximation (14) enables to calculate the approximate value of the sphere radius in cases of a tiny curvature where Pythagoras' theorem approximately rules. The explicit relation is expressed as

$$R^2 \approx \frac{(2n)^2}{n^2 + \pi^2 - s^2}\,. \tag{15}$$

The similar approximation is obtained with the help of equation (7) at the assumption of coefficients approximate equality:

$$1 \approx k_1 \approx k_2 \approx k_3\,. \tag{16}$$

Then the sphere radius is expressed as

$$R^2 \approx \frac{(\pi n)^2}{n^2 + \pi^2 - s^2}\,. \tag{17}$$

1.3 Path on the hyperbolic sphere

By the circumference of a circle concluded path s on the hyperbolic sphere is calculated with the help of the hyperbolic law of cosines.

On the hyperbolic sphere of radius R holds:

$$\cosh\frac{s}{R} = \cosh\frac{\pi}{R}\cosh\frac{n}{R}\,, \tag{18}$$

$$\cosh x = \sqrt{1 + \sinh^2 x}, \tag{19}$$

$$\frac{1}{R^2} = -\frac{1}{k_1^2\pi^2} - \frac{1}{k_2^2 n^2} + \frac{k_3^2 s^2}{k_1^2\pi^2 \times k_2^2 n^2} = \\ = \frac{-k_1^2\pi^2 - k_2^2 n^2 + k_3^2 s^2}{k_1^2\pi^2 \times k_2^2 n^2}\,. \tag{20}$$

The coefficients are expressed as

$$k_1 = \frac{\sinh\frac{\pi}{R}}{\frac{\pi}{R}}, \quad k_2 = \frac{\sinh\frac{n}{R}}{\frac{n}{R}} \quad \text{and} \quad k_3 = \frac{\sinh\frac{s}{R}}{\frac{s}{R}}\,. \tag{21}$$

They are arranged by size

$$1 \leqslant k_1 \leqslant k_2 \leqslant k_3\,. \tag{22}$$

In the case of R^2 being a positive number Pythagoras' theorem holds only exceptionally.

The next condition has to be satisfied:

$$k_3^2 s^2 \geqslant k_1^2\pi^2 + k_2^2 n^2 \quad or \quad s^2 \geqslant \frac{k_1^2}{k_3^2}\pi^2 + \frac{k_2^2}{k_3^2}n^2. \tag{23}$$

The ratios of coefficients k_1^2/k_3^2 and k_2^2/k_3^2 are according to (non)equation (22) smaller than 1 or at most equal 1, therefore we write:

$$\frac{k_1^2}{k_3^2}\pi^2 + \frac{k_2^2}{k_3^2}n^2 \leqslant \pi^2 + n^2 \leqslant s^2. \tag{24}$$

At the finite hyperbolic sphere radius R Pythagoras' theorem fails, because at non-equal coefficients (22) the square area upon hypotenuse is greater than the sum of square areas upon catheters:

$$s^2 > \pi^2 + n^2. \tag{25}$$

At $R = \infty$ and equal coefficients (22) the hyperbolic sphere transforms into the Euclidean plane and Pythagoras' theorem begins to rule again (4).

2 Fine structure constant and sphere radius

In the ground state of the Hydrogen atom the electron path around the nucleus equals the ratio of the Compton wavelength of the electron λ and the fine structure constant α. The wavelength equals the unit, so the circular path equals the inverse fine structure constant:

$$s = \alpha^{-1}. \tag{26}$$

2.1 Inverse fine structure constant on the non-Euclidean sphere and Euclidean plane

At the finite sphere radius R two possibilities are allowed according the non-equations (12) and (25).

On the elliptic sphere holds:

$$\alpha^{-2} < \pi^2 + n^2. \tag{27}$$

On the hyperbolic sphere holds:

$$\alpha^{-2} > \pi^2 + n^2. \tag{28}$$

At $R = \infty$ both non-Euclidean spheres transform into the Euclidean plane and according to the equation (4) holds:

$$\alpha^{-2} = \pi^2 + n^2. \tag{29}$$

2.2 Calculation of the theoretical inverse fine structure constant in the Euclidean plane

In the hydrogen atom the number $n = 137$ is to the inverse fine structure constant α^{-1} the closest natural number which concludes the start and end point of Bohr orbit. The number π is the total average rotation component of the length unit.

The theoretical inverse fine structure constant in the Euclidean plane is calculated with the help of the equation (29). Its value is an irrational number:

$$\alpha^{-1}{}_{\text{EUCLID}} = \sqrt{n^2 + \pi^2} \approx 137.036015720. \tag{30}$$

2.3 Calculation of the sphere radius on the atomic level

The inverse fine structure constant should be according to the equations (27) and (28) on the elliptic sphere smaller and on

the hyperbolic sphere greater than $\alpha^{-1}_{\text{EUCLID}}$.

The recommended CODATA 2012 value of the inverse fine structure constant is smaller than the theoretical value in the Euclidean plane:

$$\alpha^{-1}_{\text{CODATA}} = 137.035999074 < \alpha^{-1}_{\text{EUCLID}} \approx 137.036015720. \quad (31)$$

This implies the elliptic sphere in the Hydrogen atom.

The calculus of the radius of the elliptic sphere with the help of the equation (5) yields:

$$R = 3679 \text{ Compton wavelengths of the electron.} \quad (32)$$

The estimate of the radius of the elliptic sphere with the help of the simplified Hardy approximation (15) yields a little bit greater value:

$$R \approx \frac{2.137}{\sqrt{137^2 + \pi^2 - \alpha^{-2}_{\text{CODATA}}}} = 4057. \quad (33)$$

2.4 Estimation of the inverse fine structure constant on the macro level

Let us consider the radius of the observable universe of about 4×10^{26}m [3] as the sphere radius:

$$R \approx 2 \times 10^{38} \text{ Compton wavelengths of the electron.} \quad (34)$$

This is a huge radius. A common calculator supports the spherical law of cosines only for radius up to $\sim 10^{15}$ Compton wavelengths of the electron.

Fortunately a huge sphere radius is given by the simplified Hardy approximation (15) in the explicit relation with the inverse fine structure constant:

$$R^2 \approx \frac{(2.137)^2}{\pi^2 + 137^2 - \alpha^{-2}}, \quad (35)$$

$$\alpha^{-1} \approx \sqrt{\pi^2 + 137^2\left(1 - \frac{4}{R^2}\right)} = \sqrt{\pi^2 + 137^2\left(1 - 10^{-76}\right)} \approx$$
$$\approx \sqrt{\pi^2 + 137^2}. \quad (36)$$

If the sphere curvature on the atomic level equals the curvature of the hypothetical elliptic observable universe, the inverse fine structure constant should not significantly differ from the theoretical constant in the Euclidean plane.

3 Conclusion

If the inverse fine structure constant is a mirror of the path complexity as well as the curvature of the sphere where the path is made, its theoretical inverse value in the Euclidean plane $\alpha^{-1} = \sqrt{\pi^2 + 137^2}$ and the recommended empirical CODATA 2012 value $\alpha^{-1} = 137.035999074$ express the electron motion on the elliptic sphere of the radius of 3679 Compton wavelengths of the electron. This implies a huge curvature of the atomic world. If the sphere curvatures in the atomic and the macro-world would be the same, the inverse fine structure constant should not significantly differ from the theoretical one in the Euclidean plane.

Submitted on: October 15, 2012 / Accepted on: October 23, 2012

References

1. Binder B. Berry's Phase and Fine Structure. http://philsci-archive.pitt.edu/682/1/alfa137MN5p.pdf. Retrieved September 2012.
2. Weisstein E. W. Cosine. http://mathworld.wolfram.com/Cosine.html. Retrieved October 2012.
3. WolframAlpha. http://www.wolframalpha.com/input/?i=size+of+ universe. Retrieved November 2011.

The Poisson Equation, the Cosmological Constant and Dark Energy

Jeffrey P. Baugher

Wright State University, Department of Electrical Engineering, 311 Russ Engineering Center, 3640 Colonel Glenn Highway, Dayton, Ohio 45449, USA. E-mail: baugher.3@wright.edu

The Cosmological Constant Λ within the modified form of the Einstein Field Equation (EFE) is now thought to best represent a "dark energy" responsible for a repulsive gravitational effect, although there is no accepted argument for its magnitude or even physical presence. In this work we compare the origin of the Λ argument with the concept of unimodular gravity. A metaphysical interpretation of the Poisson equation during introduction of Λ could account for the confusion.

1 Introduction

In 1916, Einstein introduced his general theory of relativity as a geometrical theory of gravity [4] resulting in the Einstein field equation (EFE),

$$R_{\mu\nu} - \frac{1}{2}\, g_{\mu\nu}R = G_{\mu\nu} = \frac{8\pi G}{c^4}\, T_{\mu\nu}. \tag{1}$$

It has been well documented and studied that the EFE did not predict a stable static universe, as it was theorized to be at the time [3]. The equation, however, did accurately predict gravitational redshift, magnitudes of gravitational lensing and account for Mercury's precessing orbit, which the Newtonian equation could not. In order to manufacture an equation that could account for a static universe, but still be empirically accurate, it is often stated that Einstein ad hoc threw in another constant Λ which is known as the cosmological constant. This would have been placed back into the EFE with the metric $g_{\mu\nu}$ as

$$R_{\mu\nu} - \frac{1}{2}\, g_{\mu\nu}R + g_{\mu\nu}\Lambda = G_{\mu\nu}. \tag{2}$$

Once it was discovered that the universe actually appeared to be in a decelerating or coasting expansion mode, Einstein quickly removed the Λ term. Today, though, there is empirical evidence that a very small magnitude Λ exists, but some quantum field theorists estimate it as being over 120 orders of magnitude smaller than their calculations, "probably the worst theoretical prediction in the history of physics" [3]. In addition, the observed small value of Λ requires an extremely high level of arbitrary fine tuning "for no good reason" and is a "cosmologist's worst nightmare come true" [6]. This transformation from a minor but rich interest exploded (5000 papers submitted to date [10]) near the end of the past millennium due to a startling simultaneous discovery of positive acceleration from two teams [7, 8].

The source of this unforeseen positive acceleration has come to be known as "dark energy". The lack of progress in explaining the phenomena led to the creation of a Dark Energy Task Force in 2006 which stated in a report [1]:

> "Most experts believe that nothing short of a revolution in our understanding of fundamental physics will be required to achieve a full understanding of the cosmic acceleration."

This dark energy is currently expected to contribute over 73.4% [5] of the mass-energy of the universe, and there is no sound logical theory for what it is. Consider that this leaves some type of mysterious never-observed particle known as dark matter to contribute another 22.2%, leaving only 4.4% for the normal matter we are familiar with. With this in mind, we propose that it is reasonable to re-examine any argument that has lead us to our current state of physics.

2 Poisson Equation and Gauss' Theorem

The Poisson equation,

$$-\nabla^2 u = \mathbf{f}, \tag{3}$$

is well known to relate the function f as the "source" or "load" of the effect on u of the left hand side. Let us examine what this means *exactly* more in depth and what we can conclude from this tool. As an example, for a function f given on a three dimensional domain denoted by $\Omega \subset \mathbb{R}^3$ we have

$$\alpha u + \beta\, \frac{\partial u}{\partial \mathbf{n}} = \mathbf{g} \quad \text{on} \quad \partial\Omega. \tag{4}$$

This is a solution u satisfying boundary conditions on the boundary $\partial\Omega$ of Ω. α and β are constants and $\frac{\partial u}{\partial \mathbf{n}}$ represents the directional derivative in the direction normal $\mathbf{n}$ to the boundary $\partial\Omega$ which by convention points outwards. Although if $\alpha = 0$ is referred to as a Neumann boundary condition, even with $\alpha = constant$ the solution is said to only be unique up to this additive constant. Let us examine whether this statement is entirely accurate.

2.1 Graphical Meaning of Poisson Equation

Let us take the divergence of $\mathbf{g}$ so that

$$\nabla \cdot \alpha u + \nabla \cdot \beta\, \frac{\partial u}{\partial \mathbf{n}} = \nabla \cdot \mathbf{g} \tag{5}$$

and

$$0 + \nabla \cdot \beta\, \frac{\partial u}{\partial \mathbf{n}} = \nabla \cdot \mathbf{g}. \tag{6}$$

We can see that the presence of αu seems arbitrary since it has no effect. Let us examine a two dimensional slice of scalar values in $\mathbb{R}^3$ to graphically give a better understanding. In Fig. 1 we have an example of Eq. 4 using a Euclidean coordinate system.

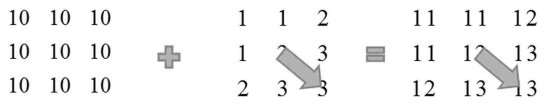

Fig. 1: Two Dimensional Scalar Field

For any derivative of Eq. 5, the constant term of course would result in no vector since there is no directional derivative from αu.

We note that this equation can also be written as

$$\alpha u - \beta \frac{-\partial u}{\partial \mathbf{n}} = \mathbf{g}, \tag{7}$$

shown in Fig. 2, which does not mathematically make a difference but can, however, introduce a question of uniqueness.

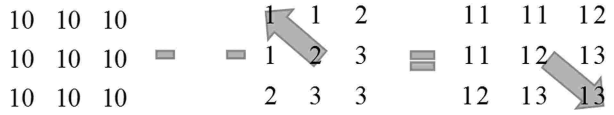

Fig. 2: Alternate Two Dimensional Scalar Field

Let us define the previous scalar field u as u_1 and a second scalar field as u_2. If ξ and γ are constants, then Eq. 8 and Fig. 3 present a dilemma. While there may be no directional derivatives from the constant term, we could also equivalently model this as orthogonal vectors with the sum of 0.

$$\xi u_2 - \gamma \frac{\partial u_2}{\partial \mathbf{n}} = \mathbf{g}_2 \tag{8}$$

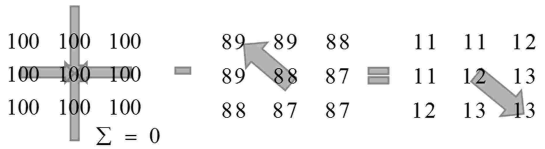

Fig. 3: Second Two Dimensional Scalar Field

From this we can see that there are no unique solutions of u for **g** from the Poisson equation, if

$$\alpha u_1 + \beta \frac{\partial u_1}{\partial \mathbf{n}} = \mathbf{g}_1 \tag{9}$$

and

$$\xi u_2 - \gamma \frac{\partial u_2}{\partial \mathbf{n}} = \mathbf{g}_2 \tag{10}$$

but also

$$\nabla \cdot (\alpha u_1 + \beta \frac{\partial u_1}{\partial \mathbf{n}}) = \nabla \cdot \mathbf{g}_1 = \nabla \cdot \mathbf{g} \tag{11}$$

and

$$\nabla \cdot (\xi u_2 - \gamma \frac{\partial u_2}{\partial \mathbf{n}}) = \nabla \cdot \mathbf{g}_2 = \nabla \cdot \mathbf{g} \tag{12}$$

if

$$\beta \frac{\partial u_1}{\partial \mathbf{n}} = -\gamma \frac{\partial u_2}{\partial \mathbf{n}}. \tag{13}$$

2.2 Gauss Theorem

Like our above illustration of the Poisson equation, a misunderstanding of Gauss' Theorem,

$$-\int_{\partial\Omega} \frac{\partial u}{\partial \mathbf{n}} = -\int_{\Omega} \nabla^2 u = \int_{\Omega} f \tag{14}$$

could also cause confusion if

$$-\int_{\partial\Omega} \beta \frac{\partial u_1}{\partial \mathbf{n}} = -\int_{\partial\Omega} \left(\xi u_2 - \gamma \frac{\partial u_2}{\partial \mathbf{n}}\right) \tag{15}$$

and

$$-\int_{\Omega} \nabla^2 \beta u_1 = -\int_{\Omega} \left(\nabla^2 \xi u_2 - \nabla \cdot \gamma \frac{\partial u_2}{\partial \mathbf{n}}\right). \tag{16}$$

Equations 15 and 16 are easily understood graphically as taking the second derivatives of the plots in Fig. 4.

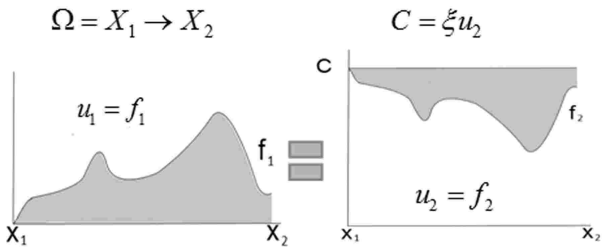

Fig. 4: Equivalent Areas From Gauss' Theorem

3 Conclusion

Although we can assume that some function **g** is causal to the appearance of a vector, does the vector appear from nothing or is it result of a change in what is already at that point? If *au* exists, what does it physically represent? Calling any field "attractive" or "repulsive" is nothing more than a metaphysical convention, i.e. does the load function cause a change in ϕ resulting in an attraction or a reduced repulsion, as in Fig. 5? From this, we can conclude that although we may possess measurements ∇u and $\nabla^2 u$, we cannot determine the nature of the scalar field u simply from the Poisson equation or Gauss' Theorem.

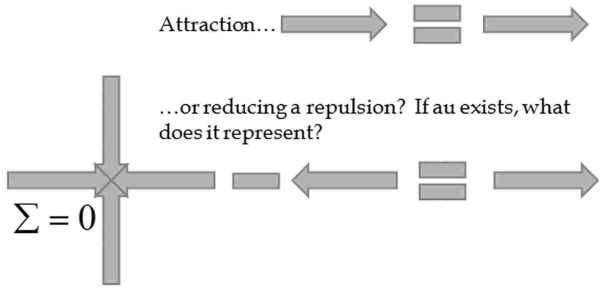

Fig. 5: Attraction or Reduced Repulsion?

4 Motivation: Cosmological Constant and General Relativity

Why is the previous figure important? Although there is a great deal of literature concerning Λ, in order to start a new perspective and to utilize the previous section, we re-examine the first known published physical meaning of the constant. In Einstein's 1917 paper *Cosmological Considerations On The General Theory of Relativity* [2] the first equation Einstein presents is the Poisson equation version of Newton's Law of Gravity

$$\nabla^2\phi = 4\pi\kappa\rho. \tag{17}$$

Citing Newtonian concerns over the limiting value of ϕ at "spatial infinity" he proposes a modification of the equation to

$$\nabla^2\phi - \lambda\phi = 4\pi\kappa\rho. \tag{18}$$

This was from an early difficulty in that the derivation required $R_{\mu\nu} = 0$ when matter or energy was not present. Due to cosmological observations though, and despite the rigor of the derivation, this requirement was eventually relaxed [4, see for relation to $G_{\mu\nu} = 0$, p. 410] allowing the introduction of a cosmological constant, even if it is not physically understood.

Setting the Poisson equation aside for the moment, it is also known that one of the interpretations of Λ or λ in Riemannian geometry is as a four dimensional constant of integration, through what is referred to as Unimodular Gravity [9]. This interpretation restricts allowable diffeomorphisms to only those preserving the four volume, but to date this has been treated as but a curious equivalent to General Relativity.

5 Introducing the Lorentz Tensor

Let us take a constant multiple of the metric $g_{\mu\nu}$ and refer to it as Ω. We do not utilize Λ or λ so as not to cause confusion and to allow us to more easily retain a difference in our understanding. Let us enforce $R_{\mu\nu} = 0$ such that

$$\Omega g_{\mu\nu} = G_{\mu\nu} + L_{\mu\nu} \tag{19}$$

where $G_{\mu\nu}$ is the Einstein tensor and $L_{\mu\nu}$ is a tensor we propose to call the "Lorentz" tensor. We shall expand on our reasoning for calling it this in subsequent papers. We can readily see that

$$G_{\mu\nu} = \Omega g_{\mu\nu} - L_{\mu\nu} \tag{20}$$

and that if $\Omega = 0$ then the Lorentz tensor is simply the negative of the Einstein tensor,

$$G_{\mu\nu} = -L_{\mu\nu}, \tag{21}$$

and should have the same important properties, i.e.

$$G_{\mu\nu;\mu} = -L_{\mu\nu;\mu}. \tag{22}$$

This of course results in

$$R_{\mu\nu} - \frac{1}{2}\, g_{\mu\nu}R = G_{\mu\nu} = \Omega g_{\mu\nu} - L_{\mu\nu}. \tag{23}$$

Note that for now cosmological models that rely on only a multiple of the metric remaining with no matter present, such as deSitter space, are not possible since $R_{\mu\nu} = 0$.

Although there are physical arguments for equating the Einstein tensor to the energy momentum tensor ($G_{\mu\nu} = \kappa T_{\mu\nu}$), and thus into analogues for Newton's Law of Gravity, we note simply in this paper that Eq. 17 is ultimately arrived at through $G_{\mu\nu}$. By the symmetry present in Eq. 23 and our arguments concerning the Poisson equation and Gauss' Theorem, our future objective is to use our understanding of Fig. 6 to obtain a rigorous derivation of Fig. 7.

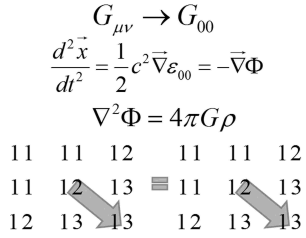

Fig. 6: Einstein Tensor to Poisson

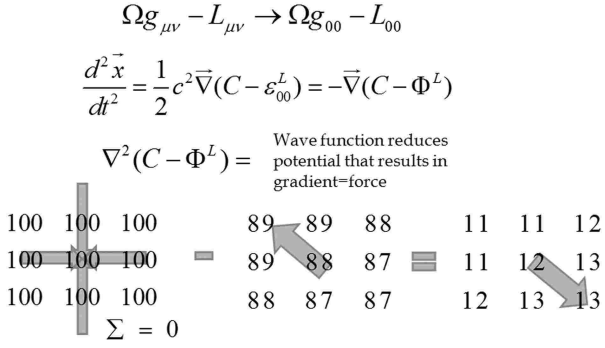

Fig. 7: Alternate EFE to Reduced Repulsive Poisson

We do this also in order to ask, should matter subject to the force represented by the vector present in Fig. 7 become

zero after traveling a certain radius from a massive body, what occurs at radii larger than this? It is our motivation to determine whether this is a plausible explanation for phenomena attributed to positive accelerating expansion.

Submitted on: October 11, 2012 / Accepted on: October 17, 2012

Acknowledgements

The author would like to thank Dr. Marian Kazimierczuk and Dmitri Rabounski.

References

1. Albrecht A., Bernstein G., Cahn R., Freedman W. L., Hewitt J., Hu W., Huth J., Kamionkowski M., Kolb E. W., Knox L. Report of the dark energy task force. arXiv: astro-ph/0609591.
2. Einstein A., Davis F.A. The principle of relativity. Dover Publications, New York, 1952.
3. Hobson M. P., Efstathiou G., Lasenby A. N., General relativity: an introduction for physicists. Cambridge University Press, Cambridge, UK ; New York, 2006.
4. Misner C. W., Thorne K. S., Wheeler J. A., Gravitation. W.H. Freeman, San Francisco, 1973.
5. NASA. Seven Year Results: Wilkinson Microwave Anisotropy Probe. 2011.
6. Padmanabhan T. Cosmological constant: the weight of the vacuum. *Physics Reports*, 2003, v. 380 (5), 235–320.
7. Perlmutter S., Aldering G., Goldhaber G., Knop R.A., Nugent P., Castro P.G., Deustua S., Fabbro S., Goobar A., Groom D.E. Measurements of Ω and Λ from 42 high-redshift supernovae. *The Astrophysical Journal*, 1999, v. 517 (2), 565–586.
8. Riess A. G., Filippenko A. V., Challis P., Clocchiatti A., Diercks A., Garnavich P. M., Gilliland R. L., Hogan C. J., Jha S., Kirshner R. P. Observational evidence from supernovae for an accelerating universe and a cosmological constant. *The Astronomical Journal*, 1998, v. 116 (3), 1009–1038.
9. Weinberg S. The cosmological constant problem. *Reviews of Modern Physics*, 1989, v. 61 (1), 1–23.
10. Author not listed. The dark side of the universe. *Economist*, 2012, v. 402 (8772), 79–81.

Magnetic Fields and Directional Spectral Emissivity in Sunspots and Faculae: Complimentary Evidence of Metallic Behavior on the Surface of the Sun

Pierre-Marie Robitaille

Department of Radiology, The Ohio State University, 395 W. 12th Ave, Columbus, Ohio 43210, USA. E-mail: robitaille.1@osu.edu

Sunspots and faculae are related phenomena and constitute regions of elevated magnetic field intensity on the surface of the Sun. These structures have been extensively studied in the visible range. In this regard, it has been recognized that the intensity contrast of faculae, relative to the photosphere, increases considerably as the line of observation moves from the center to the limb of the Sun. Such center to limb variation (CLV) suggests that the directional spectral emissivity of the faculae increases at the same time that photospheric directional emissivity decreases. Since the directional spectral emissivity of faculae increases towards the limb, these structures, along with sunspots, provide strong evidence for metallic behavior at the level of the solar surface. This further strengthens claims that the body of the Sun is not gaseous, but rather, comprised of condensed matter.

1 Introduction

In his popular work, *The Birth and Death of the Sun*, George Gamow justified the gaseous nature of the Sun as follows: "*. . . at 6000 degrees all the materials from which a furnace might be constructed, including even such refractory substances as platinum or carbon, will be not only melted but completely evaporated. No material can exist at these high temperatures in a state other than gaseous, and this is exactly what we find on the surface of the Sun, where all elements are present in vapour form*" [1, p. 4–5]. Several prominent members of the astronomy community, by utilizing similar logic, had previously laid the foundation for a gaseous Sun in the mid-1800s [2]. The contention that the Sun was too hot to be anything but gaseous would persist throughout the 20th century [3]. Conversely, experiments had long indicated that the phases of matter did not depend solely on temperature, but on factors such as external pressure, internal atomic composition, and the nature of the lattice adopted in the condensed phase. Yet, using a single justification, the possibility that certain materials might exist in liquid form within the Sun continued to be ignored. Gamow's argument [1, p. 4–5] would discount Wigner and Huntington's 1935 proposal [4] that metallic hydrogen, a material existing in the condensed phase, could be created at elevated temperatures and pressures [5–7].

2 Metallic hydrogen on the Sun

Liquid metallic hydrogen [4] is a particularly alluring substance relative to condensed solar models [5–7], especially given the observation that the Sun appears to be primarily composed of this element [8–11]. Although metallic hydrogen was first proposed nearly eighty years ago [4], it remains an elusive material in the laboratory [5]. Some claims of synthesis have received broad international acclaim [12, 13], often followed, by controversy [14–17] and slow dismissal. Others, such as claims that certain forms of metallic hydrogen can be produced in Rydberg matter, have received less attention [18].

There has recently been a new flurry of activity in the quest to produce metallic hydrogen [4] in the laboratory. In November 2011, Mikhail Eremets and Ivan Troyan published a provocative report in *Nature Materials* [19] which strongly suggested that metallic hydrogen had indeed been synthesized for the first time on the Earth. Nonetheless, given the nature of the quest for metallic hydrogen [5], it seemed crucial that more evidence be acquired [20–22]. Perhaps this time, the synthesis of metallic hydrogen will be affirmed [5].

Beyond metallic hydrogen itself, dense hydrogen could play an important role in the Sun, since the photosphere appears to be less metallic in nature than sunspots [5]. The author has advanced arguments that the photosphere adopts a layered lattice resembling graphite (a Type-1 lattice [5]), while the lattice in sunspots has more metallic character (a Type 2 lattice [5]). This is presumably due to slightly decreased inter-atomic distances within the layered lattice of sunspots. It is noteworthy that a report has recently demonstrated that dense hydrogen could adopt a graphene-like structure at 220 GPa and 300 K [23, 24]. The need for emitting a thermal spectrum provides strong motivation for considering graphite-like layered structures, which can lead to hydrogen in the metallic state, within liquid models of the Sun [5].

3 A liquid Sun

The idea that the Sun could be liquid dates back at least to the days of Gustav Kirchhoff [2] and Sir James Hopwood Jeans was its last major scientific champion [3]. Jeans was a distinguished physicist [25] and Physical Sciences Secretary of the Royal Society from 1919 to 1929 [26]. He was also Sir Arthur Eddington's principle antagonist [3]. For much of his scientific career, Jeans advanced that heavy metals such as uranium comprised the building blocks for a liquid Sun,

in opposition to Eddington's gaseous models [3]. When the Sun was determined to be principally composed of hydrogen [9–11], Jeans was left without a structural material. Eddington's gaseous Sun went on to be widely accepted by astronomy. Neither Jeans nor Eddington had anticipated the postulate that metallic hydrogen could be formed at elevated pressures [4]. For his part, Jeans abandoned the liquid model [3], apparently without sufficiently considering that the observational evidence for condensed matter might continue to mount [5–7, 27–29]. At the time, he had elucidated only fragmentary proof for a liquid state (see [3] and references therein).

Today, not a single observational line of evidence supports the idea that the Sun is gaseous, as simple temperature arguments are fallacious. Much of the scientific discussion appears centered on endowing gaseous solar models with the ability to behave as condensed matter (e.g. [30]). By dismissing the facts, the existence of the solar surface has been discounted [3], precisely because the gaseous models have no means of accounting for such a structure [29]. All structural features associated with solar activity (sunspots, faculae, prominences, flares, spicules, etc. . .) tend to be explained using magnetic fields, as the only means to impart structural features to a gaseous entity which, in reality, can support none.

In sharp contrast, observational facts point to a liquid Sun, including more than one dozen proofs for a condensed matter [5–7, 27–29]. Though the most convincing line of evidence for a liquid Sun will always remain the thermal appearance of the photospheric spectrum in the visible range [27], some may not be able to appreciate the power and sufficiency of this proof. In part, this is due to the introduction of local thermal equilibrium reasoning in solar science [30]. Local thermal equilibrium has come to cloud the requirements for producing a thermal spectrum and mask the need for condensed matter [30]. Nonetheless, the arguments which support a liquid Sun based on its thermal emission are definitive [30–33]. Thermal evidence will always remain paramount, because it points to the existence of lattice order on the surface of the Sun [31]. Nothing further is required to demonstrate the presence of condensed matter, as Kirchhoff himself indirectly understood in the mid-1800s [2]. For those who require additional illustrations, sunspots and faculae provide an interesting proving ground.

4 Directional spectral emissivity of sunspots and faculae

As key structural elements on the surface of the Sun, sunspots and faculae provide solar physicists ample opportunity for observation and discussion. In the days of Galileo and Scheiner, even the association of sunspots with the solar body was cause for extensive debate [34]. Since that time, sunspots and faculae have come to reveal much about the Sun, despite the belief that their visual appearance on the photosphere remains an optical illusion in modern solar theory [29].

4.1 Sunspots

As early as 1774, Alexander Wilson [35] noted that sunspots appeared as slight depressions relative to the solar surface. Wilson reached this conclusion based on geometry [35]. Accepted solar models currently account for the visual depression of sunspots, or "Wilson effect", using optical depth arguments (e.g [36, p. 189–190] and [37, p. 46]). Such complexity must be invoked because modern theories are built around a gaseous solar body. Since these models have long deprived the Sun of a true surface [2,29], they cannot rest upon geometrical arguments to account for the Wilson effect [35] and must have recourse to explanations based on optical depth (e.g [36, p. 189–190] and [37, 46]). Conversely, the author has argued in favor of an authentic solar surface, thereby directly challenging accepted models [29]. Hence, the Wilson effect [35], one of the oldest and simplest sunspot observation, has provided a basis for questioning the established gaseous models of the Sun.

Modern astrophysics has advanced an understanding of sunspots which, on cursory examination at least, appears to be complete. In reality, the true physical nature of these structures has remained elusive, despite our arsenal of data. Still, much has been learned about sunspots. The Wilson effect was established at the end of the 18th century [35]. Schwab discovered the eleven year sunspot cycle in 1843 [38]. In the same period, Carrington used sunspot observations and outlined the differential rotation of the Sun in great detail [39].

In 1908, George Ellery Hale discovered that sunspots are regions of powerful magnetic activity [40]. The intensity of magnetic fields at the center of sunspots has been determined to be primarily vertical and known to increase in the dark nuclei of the umbra (e.g. [37, p. 75] and [41, p. 80]). Helioseismic analysis of the Sun has revealed that sound waves travel faster within sunspots relative to the photosphere [42,43]. All of these phenomena are highly suggestive of increased density and metallicity within sunspots and have been utilized to support the idea that the Sun is condensed matter [28]. Strong magnetic fields and the science of seismology are always associated with condensed matter, not the gaseous state of solar models.

Sunspots have also been reported to have directional emissivities that increase with angle of observation, as the observer follows their movement towards the limb of the Sun [41, p. 75–77]. One of the earliest reports of increased sunspot emissivity relative to the photosphere dates back to 1875 and Samuel Langley*: "*With larger images and an improved instrument, I found that, in a complete ring of the solar surface, the photosphere, still brilliant, gave near the limb absolutely less heat than the umbra of the spots*" [44, p. 748]. Edwin Frost would soon echo Langley: "*A rather surprising result of these observations was that spots are occasionally*

*Translations from French of Langley's work [44] were executed by the author, P. M. Robitaille.

relatively warmer than the surrounding photosphere" [45, p. 143].

Should the directional emissivity of sunspots truly increase near the limb, such behavior would be highly supportive of metallic character [28]. Non-metals usually display directional spectral emissivities that tend to decrease with increasing angle of observation [46–48]. Metals often possess lower normal emissivities with respect to their directional spectral emissivities. The directional spectral emissivities of metals typically rise with increasing angle, then fall precipitously with orthogonal viewing [46–48]. Thus, a careful analysis of emissivities can provide important clues as to whether sunspots (or faculae) are behaving as metals, potentially generating strong evidence for condensed matter on the surface of the Sun.

Truly gaseous objects should be devoid of emissivities which are directionally dependent. Thus, the increased directional spectral emissivity in sunspots could only be explained with extreme difficulty using gaseous solar models and often attributed to the effect of "stray light" [41, p. 75–77]. Stray light arguments have played an important role in the modern dismissal of increased emissivity in sunspots towards the solar limb. Thus, despite 100 years of study, the exact directional emissivity within these objects remains an unresolved issue in solar physics. The same cannot be said of facular directional spectral emissivity.

4.2 Faculae

The directional spectral emissivity contrast of faculae, with respect to the photosphere, has long been known. George Ellery Hale wrote, relative to the emissivity of the faculae: "*The bright faculae, which rise above the photosphere, are conspicuous when near the edge of the Sun, but practically invisible when they happen to lie near the center of the disk...*" [49, p. 85–86]. Hale later re-emphasized the changing emissivity of the faculae as a function of position on the solar disk: "*Mention has already been made of the faculae, which are simply regions in the photosphere that rise above the ordinary level. Near the edge of the Sun, their summits lie above the lower and denser part of that absorbing atmosphere which so greatly reduced the Sun's light near the limb, and in this region the faculae may be seen visibly. At times they may be traced to considerable distances from the limb, but as a rule they are inconspicuous or wholly invisible towards the central part of the solar disk*" [49, p. 90].

In 1961, Rogerson presented an elegant summary of the increase in facular directional emissivity observed near the solar limb [50]. This work was complemented with theory and a few photographs [50]. Rogerson noted that the contrast variation between the faculae and photosphere increased to a maximum of about 64% near the very limb of the Sun [50]. Today, the center to limb variation (CLV) of facular emissivity is widely accepted and studied [51–54], as has the grouping of faculae with sunspots (e.g. [55, p. 42–43] and [56, p. 248–249]), and the identification of faculae as regions of intense magnetic activity [57–59].

The association of bright faculae with sunspots can be traced at least to the middle of the 19th century. According to de la Rue and his team, in 1865: "*It would thus appear as if the luminous matter being thrown up into a region of greater absolute velocity of rotation fell behind to the left; and we have thus reason to suppose that the faculous matter which accompanies a spot is abstracted from that very portion of the Sun's surface which contains the spot, and which has in this manner been robbed of its luminosity*" [60]. This direct association of sunspot and facular matter has recently been re-emphasized as a result of studying large flares on the solar surface [61].

While faculae display CLV with respect to their spectral emissivity, their emissivity contrast remains highly associated with the magnetogram signal [59]. Facular contrast, after increasing to a maximum near $\mu = 0.2$ (where $\mu = \cos\theta$ and θ is the heliocentric angle between the pixel of interest and direction of the Earth; r, the distance from the disk center, is given by $r = R\sin\theta$, if R represents the solar radius) has been observed to drop rapidly when moving even closer towards the limb [52]. This finding [52] appears to be in agreement with Spruit's "*hot wall*" model of facular emissivity [62, 63].

Spruit's "*hot wall*" model stated that faculae appeared darker when viewed directly from above because very little of the "*hot wall*" was visible. As the faculae moved towards the limb, the "*hot wall*" became increasingly visible and, hence, the structures appeared bright. With increasing distance towards the limb, the "*hot wall*" once again fell out of the line of sight, being obscured by the trailing wall, and the faculae once again appeared darker (see [53] for additional detail). Others have reported that facular contrast continues to increase towards the limb (e.g. [51]). This behavior would be more consistent with the "*hot cloud*" model [50, 64, 65]wherein the faculae are viewed as floating above the photosphere [53]. Today, Spruit's "*hot wall*" model has gained almost universal acceptance, as more in accordance with observation (e.g. [66, 67]).

Alternatively, it is herein proposed that the directional spectral emissivity observed in faculae constitutes one of the most elegant proofs that the Sun is comprised of condensed matter. The reasoning remains that advanced in section 3.1 (see also [28]), with the important distinction that the directional spectral emissivity changes in faculae, unlike sunspots, are uncontested [51–54, 57–59, 66, 67]. Moreover, the observation that directional spectral emissivity contrast in faculae increase towards the limb, before rapidly subsiding at the very edge of the Sun [52], strongly supports metallic behavior in these structures [28, 46–48].

On the Earth, the existence of directional spectral emissivity in condensed matter has been established [46–48, 68]. Materials display emissivities which always manifest their

atomic nature and structure, in addition to the temperature of observation [46–48, 68]. Every material possesses a unique signature and this constitutes a powerful lesson from the study of condensed matter [46–48, 68].

The idea that faculae are condensed matter based on directional emissivities also gains support from the realization that these objects, like sunspots, are regions of intense magnetic activity [57–59]. The ideal means of accounting for this activity remains the invocation of conduction bands. A solar body which is comprised of liquid metallic hydrogen and adopts a layered graphite-like lattice presents a wonderful material to account both for the directional spectral emissivities of faculae and the associated high magnetic field [5, 28]. While condensed matter can easily support such fields, there remains no evidence on the Earth that gases, in isolation, can generate powerful magnetic fields. While it is true that gaseous plasmas respond to the presence of magnetic fields, they certainly do not possess the required structure to create such phenomena.

5 Conclusion

Despite the wide acceptance of Spruit's "*hot wall*" model of facular emissivity [62] numerous problems exist with such approaches.

First, modern models of solar emissivity are fundamentally dependent on elemental and ionic opacities within the Sun. However, the solar spectrum cannot be generated using the sum of individual opacities. The author has designated solar opacity as the Achilles' heel of the gaseous solar models [30]. It is not reasonable to account for solar emission with phenomena which cannot explain the simple emissivity found on the Earth within graphite [30].

Second, a discussion of facular emissivity often focuses on local thermal equilibrium (LTE) arguments (e.g. [66]) and such arguments are not applicable to the Sun [30]. The Sun operates well outside the confines of local thermal equilibrium and Milne's argument in support of such a regimen [69–72] leads to conduction, not equilibrium [30].

Third, the assignment of temperatures, based on emissivities on the solar surface, constitutes a direct violation of the principles associated with thermal emission [30–33], as has been highlighted by Max Planck himself [73, §101] and discussed in detail [74].

Finally, the idea that a fully gaseous object can support structure remains contrary to the known principles of physics. Objects such as "*walls*", even when only considering emissivity, require condensed matter. They cannot be mimicked by gases with densities approaching that of the best vacuums achievable on the Earth [27].

In modern solar theory, sunspots are thought to be dark, as the magnetic fields they contain prevent hot gases from rising from the interior of the Sun (e.g. [75]). Conversely, the brightness of faculae are explained when magnetic fields dilute the solar material beneath them and causes the light to escape more easily. These explanations constitute stark contrasts with one another, while at the same time discounting much of what is known on the Earth relative to thermal emissivity. The fact remains that gases are unable to emit photons in a directionally dependent manner. Astrophysical explanations relative to the causes of directional emissivity, as related to photospheric limb darkening, solar granulations, sunspots, and faculae, with their reliance on "optical depth" and "solar opacities", remain at a serious disadvantage, relative to solar models based on condensed matter [27–30].

Irrespective of the mathematical elegance associated with modern solar models, there is no observational support that the body of the Sun is a gas. Given the nature of the solar spectrum, seismic activity, and the presence of structural entities such as sunspots, prominences, and faculae, modern theory must constantly resort to mathematical arguments, or the presence of magnetic fields, in order to endow a gaseous Sun with the properties of condensed matter [8–10]. In reality, while the corona displays features consistent with gaseous plasma, the photosphere, with its sunspots, faculae, and eruptive prominences, strongly manifests the condensed nature of the solar body. The idea that solar temperatures forbid the formation of condensed matter in the Sun ignores the reality that the phases of matter are not solely determined by temperature, but are a manifestation of many factors, including pressure of formation and the internal physical properties of materials [5–7].

Currently, numerous lines of evidence strongly support the condensed nature of the Sun. These include:

1) the continuous nature of the thermal spectrum [6, 27–30],
2) photospheric limb darkening [27, 28],
3) the absence of solar collapse [5, 6, 27],
4) a solar density (1.4 g/cm^3) consistent with a hydrogen lattice [6, 27],
5) the presence of seismic activity [6, 27],
6) the behavior of mass displacement on the solar surface [6, 27],
7) the chromosphere and critical opalescence [27],
8) the existence of solar oblateness [6, 27],
9) the extensive surface activity [6, 27, 28],
10) the orthogonal nature of photospheric/coronal flows [27],
11) the ability to image the solar surface [6, 27–29],
12) the presence of a powerful solar dynamo [27],
13) the nature and behavior of sunspots, including the Wilson effect [27, 28], and
14) the structure and dynamic evolution of solar granulation [28].

Each of these phenomena can be readily incorporated into a condensed model of the Sun. Conversely, gases can neither support nor act as structural entities. A striking example relative to thermal emission and the solar opacity problem in gaseous models has been addressed in detail [30].

In this work, a fifteenth line of evidence for the condensed nature of the Sun is presented:

15) the directional spectral emissivity of faculae. Emissivity fundamentally reflects a "Planckian proof" or a "thermal proof" for condensed matter. Along with 1) the thermal appearence of the solar spectrum, 2) the limb darkening of the photosphere, 3) the directional spectral emissivity of sunspots, and 4) the directional spectral emissivity of granulations [28], the emissivity of faculae constitutes one of the most powerful lines of evidence that the Sun is condensed matter. It therefore represents the fifth thermal proof for condensed matter on the surface of the Sun.

It remains highly likely that the Planckian proofs constitute direct physical evidence for a solar lattice [31]. Through the study of directional spectral emissivity, they argue for metallicity both within sunspots and faculae. Such metallicity represents a manifestation of the lattice and the conduction bands which it supports. The Planckian proofs also remind us of the need to properly address and understand complex emission mechanisms. Driven by a desire to better comprehend the solar spectrum, perhaps someday, the physics community, at last, will link thermal emission to a unique physical process as the author has suggested [31–33]. In so doing, condensed matter and theoretical physicists will finally conclude the work initiated, but left unfinished, by Max Planck [73].

Dedication

The work is dedicated to Professor Manuel Tzagournis, Senior Vice-President for Health Science and Dean of the College of Medicine (Emeritus) at The Ohio State University for the faith he placed in realizing the dreams and hopes of a young assistant professor.

Submitted on: November 01, 2013 / Accepted on: November 03, 2013
First published in online: November 04, 2012

References

1. Gamow G. The Birth and Death of the Sun. A Mentor Book: The New American Library, New York, N.Y., 1952.
2. Robitaille P.M. A thermodynamic history of the solar constitution — I: The journey to a gaseous sun. *Progr. Phys.*, 2011, v. 3, 3–25.
3. Robitaille P.M. A thermodynamic history of the solar constitution — II: The theory of a gaseous sun and Jeans' failed liquid alternative. *Progr. Phys.*, 2011, v. 3, 41–59.
4. Wigner E. and Huntington H.B. On the possibility of a metallic modification of hydrogen. *J. Chem. Phys.*, 1935, v.3, 764–770.
5. Robitaille P.M. Liquid metallic hydrogen: A building block for the liquid Sun. *Progr. Phys.*, 2011, v. 3, 60–74.
6. Robitaille P.M. A high temperature liquid plasma model of the Sun. *Progr. Phys.*, 2007, v. 1, 70–81 (also in arXiv: astro-ph/0410075).
7. Robitaille P.M. The Sun as a high energy/high density liquid metallic hydrogen plasma. *The 33rd IEEE International Conference on Plasma Science*, June 4–8, 2006, Traverse City, Michigan, p. 461, DOI:10.1109/PLASMA.2006.1707334.
8. Payne C.H. The relative abundances of the elements. Stellar Atmospheres. Harvard Observatory Monograph No. 1, Harvard University Press, Cambridge, MA, 1925, Chapter 13 (reprinted in part in Lang K.R. and Gingerich O. A source book in astronomy and astrophysics, 1900–1975, Harvard University Press, Cambridge, MA, 1979, p.245–248).
9. Unsöld A. Über die Struktur der Fraunhofersehen Linien und die quantitative Spektralanalyse der Sonnenatmosphäre. *Zeitschrift für Physik*, 1928, v. 46, 765–781.
10. Russell H.N. On the composition of the Sun's atmosphere. *Astrophys. J.*, 1929, v.70, 11–82.
11. Grevesse N. and Sauval A.J. Standard solar composition. *Space Science Reviews*, 1998, v. 85, 161–174.
12. Weir S.T., Mitchell A.C. and Nellis W.J. Metallization of fluid molecular hydrogen at 140 GPa (1.4 Mbar). *Phys. Rev. Letters*, 1996, v. 76(11), 1860–1863.
13. Mao H.K. and Hemley R.J. Optical studies of hydrogen above 200 Gigapascals: Evidence for metallization by band overlap. *Science*, 1989, v. 244, 1462–1464.
14. Silvera I.F. Evidence for band overlap metallization of hydrogen. *Science*, 1990, v. 247(4944), 863.
15. Mao H.K. and Hemley R.J. Evidence for band overlap metallization of hydrogen — Response. *Science*, 1990, v. 247(4944), 863–864.
16. Eggert J.H., Moshary F., Evans W.J., Lorenzana H.E., Goettel K.A. and Silvera I.F. Absorption and reflectance in hydrogen up to 230 GPa: Implications for metallization. *Phys. Rev. Letters*, 1991, v. 66, 193–196.
17. Besson J.M. Comment on "Metallization of fluid molecular hydrogen at 140 GPa (1.4 Mbar)". *Phys. Rev. Letters*, 1997, v. 78(26), 5026.
18. Holmlid L. Sub-nanometer distances and cluster shapes in dense hydrogen and in higher levels of hydrogen Rydberg matter by phase-delay spectroscopy. *J. Nanopart. Res.*, 2011, v. 13, 5535–5546.
19. Eremets M.I. and Troyan I.A. Conductive dense hydrogen. *Nature Materials*, 2011, v. 10, 927–931.
20. Research Highlights. Hydrogen made metallic. *Nature*, 2011, v. 479, 448.
21. Jephcoat A.P. High-pressure physics: Testing one's metal. *Nature Materials*, 2011, v. 10, 904–905.
22. Cartwright J. Chemists claim metallic hydrogen creation first. *Chemistry World*, November 14, 2011.
23. Howie R.T., Guillaume C.L., Scheler T., Goncharov A.F. and Gregoryanz E. Mixed molecular and atomic phase of dense hydrogen. *Phys. Rev. Letters*, 2012, v. 108(12), 125501 (5 pages).
24. Cartwright J. Hydrogen that mimics graphene. *Chemistry World*, April 2, 2012.
25. Milne E.A. Sir James Jeans — A Biography. Cambridge University Press, Cambridge, 1952.
26. http://royalsociety.org/about-us/governance/officers/ (accessed 10/15/2012)
27. Robitaille P.M. The solar photosphere: Evidence for condensed matter. *Progr. Phys.*, 2006, v. 2, 17–21.
28. Robitaille P.-M. On Solar granulations, limb darkening, and sunspots: Brief insights in remembrance of Father Angelo Secchi. *Progr. Phys.*, 2011, v. 3, 79–88.
29. Robitaille P.-M. On the presence of a distinct Solar surface: A reply to Hervé Faye. *Progr. Phys.*, 2011, v. 3, 75–78.

30. Robitaille P.M. Stellar opacity: The Achilles heel of the gaseous Sun. *Progr. Phys.*, 2011, v. 3, 93–99.
31. Robitaille P.M.L. On the validity of Kirchhoff's law of thermal emission. *IEEE Trans. Plasma Sci.*, 2003, v. 31(6), 1263–1267.
32. Robitaille P.-M. Blackbody radiation and the carbon particle. *Progr. Phys.*, 2008, v. 3, 36–55.
33. Robitaille P.-M. Kirchhoff's law of thermal emission: 150 years. *Progr. Phys.*, 2009, v. 4, 3–13.
34. Galilei G. and Scheiner C. On Sunspots. Translated by E. Reeves and A.V. Helden, University of Chicago Press, Chicago, 2010.
35. Wilson A. Observations on the solar spots. *Phil. Trans. Roy. Soc.*, 1774, v. 64, 1–30.
36. Tandberg-Hanssen E. Solar Activity. Blaisdell Publishing Co., Waltham, M.A., 1967.
37. Thomas J.H. and Weiss N.O. Sunspots and Starspots. Cambridge University Press, Cambridge, U.K., 2008.
38. Schwab H. Sonnenbeobachtungen im Jahre 1843. Von Herrn Hofrath Schwabe in Dessau. *Astronomische Nachrichten*, 1844, v. 21, 233–236.
39. Carrington R.C. Observations on the Spots of the Sun, from November, 9, 1853, to March 24, 1861, Made at Redhill. Williams and Norgate, London, U.K., 1863.
40. Hale G.E. On the probable existence of a magnetic field in Sun-spots. *Astrophys. J.*, 1908, v. 28, 315–343.
41. Sobotka M. Fine structure in sunspots. In: *Motions in the solar atmosphere* (A. Hanslmeier and M. Messerotti, eds.), Astrophysics and Space Science Library, v. 239, Kluwer Academic Publishers, Dordrecht, 1999, p.71–97.
42. Moradi H. and Cally P.S. Time-distance modeling in a simulated sunspot. *Solar Physics*, 2008, v. 251, 309–327.
43. Ilonidis S. and Zhao J. Determining absorption, emissivity reduction, and local suppression coeffcients inside sunspots. *Solar Physics*, 2011, v. 268, 377–388.
44. Langley S. Sur la température des diverses régions du soleil. Les noyaux noirs des taches. *Comptes Rendus*, 1875, v. 80, 746–749.
45. Frost E.B. Observations on the thermal absorption in the solar atmosphere made at Potsdam. *Astronnomische Nachrichten*, 1892, v. 130 (3105–3106), 129–146.
46. Modest M.F. Radiative heat transfer. McGraw-Hill, New York, 1993, p.92–108.
47. Thirumaleshwar M. Fundamentals of Heat and Mass Transfer. Dorling Kindersley, Dehli, 2009, p. 652.
48. Incropera F.P., DeWitt D.P., Bergman T.L., Lavine A.S. Fundamentals of Heat and Mass Transfer, 6th Edition. JohnWiley & Sons, Hoboken, NJ, 2007.
49. Hale G.E. The study of stellar evolution: An account of some recent methods of astrophysical research, The decennial publications of the University of Chicago — Second Series, Vol. X. University of Chicago Press, Chicago, 1908.
50. Rogerson J.B. On photospheric faculae. *Astrophys. J.*, 1961, v. 134, 331–338.
51. Chapman G.A. and Klabunde D.P. Measurements of the limb darkening of faculae near the solar limb. *Astrophys. J.*, 1982, v. 261, 387–395.
52. Libbrecht K.G., Kuhn J.R. On the facular contrast near the solar limb. *Astrophys. J.*, 1985, v. 299, 1047–1050.
53. Lawrence J.K. and Chapman G.A. Photometric observations of facular contrasts near the solar limb. *Astrophys. J.*, 1988, v. 335, 996–1004.
54. Berger T.E., van der Voort L.R., Löfdahl M. Contrast analysis of solar faculae and magnetic bright points. *Astrophys. J.*, 2007, v. 661, 1272–1288.
55. Wilson P.R. Solar and stellar activity cycles. Cambridge University Press, Cambridge, U.K., 1994.
56. Bray R.J. and Loughhead R.E. Sunspots. Chapman and Hall Ltd., London, U.K., 1964.
57. Chapman G.A. Facular line profiles and facular models. *Astrophys. J. Supp. Ser.*, 1977, v. 33, 35–54.
58. Tarbell T.D. and Title A.M. Measurements of magnetic fluxes and field strengths in the photospheric network. *Solar Physics*, 1977, v. 52, 13–25
59. Ortiz A., Solanki S.K., Domingo V., Fligge M. and Sanahuja B. On the intensity contrast of solar photospheric faculae and network elements. *Astron. & Astrophys.*, 2002, v. 388, 1036–1047.
60. de la Rue W., Stewart B. and Loewy B. Researches on solar physics — Series II: On the behaviour of sun-spots with regard to increase and diminution (abstract). *Proc. Roy. Soc. London*, 1865, v. 14, 59–63.
61. Wang H., Deng N. and Liu C. Rapid transition of uncombed penumbrae to faculae during large flares. *Astrophys. J.*, 2012, v. 748(2), 76.
62. Spruit H.C. Pressure equilibrium and energy balance of small photospheric fluxtubes. *Solar Physics*, 1976, v. 50, 269–295.
63. Walton S.R. Flux tube models of solar plages. *Astrophys. J.*, 1987, v. 312, 909–929.
64. Kononovich E.V. A unified interpretation of photospheric and facular granules. *Soviet Astronomy Letters*, 1979, v. 5, 50–52.
65. Schatten K.H., Mayr H.G., Omidvar K., Maier E. A hillock and cloud model for faculae. *Astrophys. J.*, 1986, v. 311, 460-473.
66. Unruh Y.C., Solanki S.K. and Fligge M. The spectral dependence o facular contrast and solar irradiance variations. *Astron. & Astrophys.*, 1999, v. 345, 635–642.
67. Keller C.U., Schüssler M., Vögler A. and Zakharov V. On the origin of solar faculae. *Astrophys. J. Letters*, 2004, v. 607, L59–L62.
68. Touloukian Y.S. and Ho C.Y. Thermophysical Properties of Matter (vols. 1). Plenum, New York, 1970.
69. Milne E.A. Selective radiation-pressure and the structure of a stellar atmosphere. *Mon. Not. Roy. Astron. Soc.*, 1927, v.87, 697–708.
70. Milne E.A. The effect of collisions on monochromatic radiative equilibrium. *Mon. Not. Roy. Astron. Soc.*, 1928, v. 88, 493–502.
71. Milne E.A. Bakerian Lecture: The structure and opacity of a stellar atmosphere. *Phil. Trans. Roy. Soc. London*, 1929, v. 228, 421–461.
72. Milne E.A. Thermodynamics of the stars. *Handbuch der Astrophysik*, 1930, v. 3, Part 1, 65–255.
73. Planck M. The Theory of Heat Radiation. P. Blakiston's Son & Co., Philadelphia, PA, 1914.
74. Robitaille P.M. On the temperature of the photosphere: Energy partition in the Sun. *Progr. Phys.*, 2011, v. 3, 89–92.
75. http://atst.nso.edu/files/press/ATST_book.pdf (accessed on 10/25/2012 — see Page 8).

Evolution of Stellar Objects According to J. Wheeler's Geometrodynamic Concept

Anatoly V. Belyakov
E-mail: belyakov.lih@gmail.com

The proposed model is based on J. Wheeler's geometrodynamic concept, in which space continuum is considered as a topologically non-unitary coherent surface admitting the existence of transitions of the input-output kind between distant regions of the space in an additional dimension. The existence of closed structures (macrocontours) formed at the expense of interbalance of gravitational, electric, magnetic and inertial forces has been substantiated. It is such macrocontours that have been demonstrated to form — independently of their material basis — the essential structure of stellar objects (SO) and to determine the position of these objects on the Hertzsprung-Russell diagram. Models of the characteristic types of stellar objects: stars and compact bodies emerging in the end of stellar evolution — have been presented, and their standard parameters at different stages of evolution have been calculated. The existence of the Hertzsprung-Russell diagram has been substantiated, and its computational analogue has been given. Parallels between stellar and microcosmic objects are drawn.

Recognizing the Seeker, Nature
itself will come to meet him.

Rockwell Kent

1 Introduction

Wheeler's geometrodynamic concept, in which microparticles are considered as vortical oscillating deformations on a non-unitary coherent surface, was earlier used by the author to construct model objects of the microcosm [1, 2]. Those works substantiated the existence of closed structures (contours), determining the properties of microparticles. At the same time, the idea about transitions between distant regions of space in the form of Wheeler's "wormholes" can be extended to the scale of macrocosm, and some contemporary astrophysical theories has already made use of it [4]. In this paper, the existence of closed contours is substantiated at the cosmological scale, and grounds are given that they make the basis of stellar objects (SO).

The work does not consider the nature of the cosmological medium that forms stellar bodies, nor it does the nature of mass/charge carriers, force interactions etc., or various physical *manifestations* of the evolutionary behavior of stellar objects. These tasks are a subject of specific disciplines.

The model presented in the paper has an outline, illustrative character and suggests a new look at the problem. For the model, the only important thing is the *existence* of the aforementioned *entities*, forming certain types of stellar structures and determining their evolution. The work does use specific SO terms, but only schematic SO models are considered, with their evolution depending only on a few parameters reflecting the most important features of the real objects.

The SO models used here are based on the balance between main interactions: electrical, magnetic, gravitational and inertial — with no additional coefficients introduced. The analysis gives good qualitative results and, in a number of cases, plausible quantitative parameters for the statistically averaged (typical) stellar objects.

2 Initial premises

As was shown earlier [1], from the purely mechanistic point of view the so-called *charge* only manifests the degree of the nonequilibrium state of physical vacuum; it is proportional to the momentum of physical vacuum in its motion along the contour of the vortical current tube. Respectively, the *spin* is proportional to the angular momentum of the physical vacuum with respect to the longitudinal axis of the contour, while the *magnetic interaction* of the conductors is analogous to the forces acting among the current tubes.

It is given that the elementary unit of such tubes is a unit with the radius and mass close to those of a classical electron (r_e and m_e).

It should be noted that in [1, 2] the expressions for the electrical and magnetic forces are written in a "Coulombless" form, with charge replaced by electron limiting momentum. In this case, the electrical and magnetic constants (ε_0 and μ_0) are expressed as follows:

$$\varepsilon_0 = \frac{m_e}{r_e} = 3.33 \times 10^{-16} \text{ kg/m}, \tag{1}$$

$$\mu_0 = \frac{1}{\varepsilon_0 c^2} = 0.0344 \text{ N}^{-1}. \tag{2}$$

The electrical constant here is, in fact, the linear density of the vortex tube, with the mass:

$$m = \varepsilon_0 \, l \,, \tag{3}$$

where l is the length of the vortex tube (thread) or contour.

To combine the interactions, let us express them in a dimensionless form with the common force dimension factor $\frac{1}{\mu_0}$. Taking into account (1) and (2),

$$F_e = \frac{1}{\mu_0}\left(\frac{r_e}{r_0}\right)^2 z_{e_1} z_{e_2}, \tag{4}$$

$$F_m = \frac{1}{\mu_0}\frac{l}{2\pi r_0}\frac{r_e^2}{(c\times[\mathrm{sec}])^2} z_{e_1} z_{e_2}, \tag{5}$$

$$F_g = \frac{1}{\mu_0}\frac{1}{f}\left(\frac{r_e}{r_0}\right)^2 z_{g_1} z_{g_2}, \tag{6}$$

$$F_i = \frac{1}{\mu_0}\frac{r_e}{r_0}\left(\frac{\mathrm{v}_0}{c}\right)^2 z_g, \tag{7}$$

where v_0, r_0, z_e, z_g, f are the rotary velocity and rotary radius or distance between the vortex tubes, the relative values of charge and mass in the parameters of electron charge and mass and the ratio of electrical-to-gravitational forces, which, under the given conditions, is expressed as follows:

$$f = \frac{c^2}{\varepsilon_0 \gamma} = 4.16\times 10^{42}, \tag{8}$$

where γ is the gravitational constant.

The *balance of electrical and magnetic forces* $F_e = F_m$ gives a geometrical mean, a characteristic linear parameter that is independent of the direction of the vortex tubes and the number of charges

$$R_\odot = \sqrt{r_0\, l} = \sqrt{2\pi}\, c\times[\mathrm{sec}] = 7.52\times 10^8 \text{ m}, \tag{9}$$

a magnitude close to the Sun radius and the sizes of typical stars.

The *balance of magnetic and gravitational forces* $F_m = F_g$ also results in a geometrical mean:

$$\sqrt{r_0\, l} = \sqrt{\frac{z_{g_1} z_{g_2}}{z_{e_1} z_{e_2}}}\sqrt{\frac{2\pi}{f}}\, c\times[\mathrm{sec}] = \sqrt{\frac{\varepsilon}{f}}\, R_\odot, \tag{10}$$

where the ratio of the products $\varepsilon = z_{g_1} z_{g_2}/z_{e_1} z_{e_2}$ is an *evolutionary parameter*, which characterizes the state of the medium and its changes, as the mass carriers become predominant over the electrical ones and, as a matter of fact, shows how the material medium differs from vacuum.

In the general case, expression (10) gives a family of lengthy contours, consisting of contra-directional closed vortex tubes (*m_g-contours*). The evolutionary parameter ε proportionally increases the mass of the vortex tube for the m_g-element:

$$m = \varepsilon\varepsilon_0\, l. \tag{11}$$

The vortex tubes can consist, in their turn, of a number of parallel vortex threads, whose stability is ensured by the *balance of magnetic and inertial forces* ($F_m = F_i$; *m_i-zones*). As follows from this balance,

$$\mathrm{v}_{0i} = \sqrt{\frac{z_{e_1} z_{e_2}}{z_g}}\sqrt{\frac{r_e\, l}{2\pi}}\times[\mathrm{sec}^{-1}]. \tag{12}$$

Unidirectional vortex threads of the length l rotate, with the rotary velocity v_{0i}, about the longitudinal axis along an orbit of indeterminate radius. When they are filled with the chains of single charges, having the mass of an electron, and their number $z_e = z_g = l/r_e$ (or when the tubes consist of single vortex threads in the quantity of l/r_e), we get the following equation:

$$\mathrm{v}_{0i} = \frac{l}{\sqrt{2\pi}}\times[\mathrm{sec}^{-1}]. \tag{13}$$

The *balance of gravitational and inertial (centrifugal) forces* $F_g = F_i$ gives a *virial*, from which one can derive the maximal gravitational mass of the object, satisfying condition (9):

$$M_m = \frac{R_\odot c^2}{\gamma} = f R_\odot \varepsilon_0 = 1.012\times 10^{36} \text{ kg}. \tag{14}$$

3 Structurizations of the primary medium and parameters of stellar objects

Now let us consider objects in which more than one pair of forces is balanced.

Let us assume that an initially unstructured maximal mass evolves and becomes more complex — through the emergence of m_i-zones, consisting of single elements of the length l_i and mass m_i. As follows from the constancy of μ_0 in the general case,

$$\frac{1}{\mu_0} = \varepsilon_0 c^2 = \frac{m_i \mathrm{v}_{i_0}^2}{r_i} \tag{15}$$

where $m_i = \varepsilon_0 l_i$ is the mass of a vortex m_i-element. From (13) and (15), one can obtain, having in mind (9), the ratio for its geometrical parameters:

$$\frac{l_i^3}{r_i} = R_\odot^2. \tag{16}$$

Driven by gravitation, the single tubes (threads) will combine into a local structure, the mass of which can also be calculated from the virial:

$$M_i = \frac{r_i \mathrm{v}_{i_0}^2}{\gamma}. \tag{17}$$

Let the object contain z_i local zones; then its mass will be $M_0 = z_i M_i$. Let us introduce a dimensionless parameter

$M = M_0/M_m$. Then, making some transformations, one can eventually obtain uniform equations for all the parameters of the evolving objects with an arbitrary relative mass M:

number of local zones

$$z_i = \frac{1}{M^{1/4}}\,, \tag{18}$$

zone radius

$$r_i = M^{3/4} R_\odot\,, \tag{19}$$

length of the vortex tube (thread)

$$l_i = M^{1/4} R_\odot\,, \tag{20}$$

rotary velocity in the zone

$$\mathrm{v}_{0i} = M^{1/4}\, c, \tag{21}$$

number of single vortex threads in the zone

$$n = \frac{M_i}{m_i} = f M\,, \tag{22}$$

and, having in mind (10), one can take $n = \varepsilon$.

Thus, as its mass decreases, the *object simultaneously becomes more and more complex*, getting subtly structured with m_i-zones.

Let us assume that the initial state of SO is a rotating disk, which can further develop into larger structures (m_g-contours) of the size $R_0 \times d_0$, where the contour length is $R_0 = l$ and diameter is $d_0 = r_0$. With these designations, equation (10) will look as follows:

$$\sqrt{d_0\, R_0} = \sqrt{\frac{\varepsilon}{f}}\, R_\odot\,. \tag{23}$$

Let us accept, quite schematically and roughly, that m_g-contours in the disk are oriented radially-spirally and are pulled in towards the center by the radial components of the gravitational forces. These forces are approximately equal to $(d_0/R_0)F_g$. Then, from the balance of centrifugal and gravitational forces,

$$\mathrm{v}_0 = \sqrt{\frac{d_0}{R_0}}\,\sqrt{\frac{\gamma m}{R_0}}\,, \tag{24}$$

where m and R_0 are the m_g-contour mass and the averaged disk radius respectively.

Let us define the number of m_g-contours as

$$z_0 = \frac{R_0}{d_0}\,. \tag{25}$$

With equation (11) in mind, the total mass of the object will amount to

$$M M_m = z_0\, m = z_0\, \varepsilon \varepsilon_0\, R_0. \tag{26}$$

Taking into account equations (8), (9), (23–26) and making some transformations, we can find parameters of the structured disk:

$$R_0 = M^{1/3}\, R_\odot\,, \tag{27}$$

$$z_0 = \frac{f M^{2/3}}{\varepsilon}\,, \tag{28}$$

$$\mathrm{v}_0 = \frac{\varepsilon c}{f M^{1/3}}\,. \tag{29}$$

The parameters found are averaged when the disk structural elements are tightly packed, and they determine the core of the object. Let us define the object boundaries — under the condition that, if the system of m_g-contours is rotating as a rigid disk, the rotary velocity of contours at the periphery must not exceed the speed of light. In this case, the maximal radius of the disk will be:

$$R_m = \frac{R_0\, c}{\mathrm{v}_0} = z_0\, R_\odot\,. \tag{30}$$

Let us further assume — within the framework of our simplified model — that the mass of the object is concentrated either in the center (the *state of core*) or at the periphery (the *state of outer layer*). Obeying the angular momentum conservation law, velocity at the periphery cannot be higher than:

$$\mathrm{v}_m = \frac{\mathrm{v}_0 R_0}{R_m} = \frac{\mathrm{v}_0^2}{c}\,. \tag{31}$$

Let the periods of core and outer layer rotation be expressed as $\tau_0 = R_0/\mathrm{v}_0$ and $\tau_m = R_m/\mathrm{v}_m$ respectively (the duration of the inner and outer cycles).

Having in mind (27–31) and taking into account that $\sqrt{2\pi} = 2.51$, we obtain

$$\tau_0 = 2.51 M^{2/3}\frac{f}{\varepsilon}, \tag{32}$$

$$\tau_m = 2.51 M^{4/3}\left(\frac{f}{\varepsilon}\right)^3\,. \tag{33}$$

Indeed, star cores rotate much faster that their outer layers [5]. As the medium condenses and becomes more and more different from vacuum, the evolutionary parameter ε grows. There are at least two characteristic values of this parameter satisfying the following conditions:

1. The number of m_g-elements z_0 is equal to the number of m_i-structures z_i, which should correspond to the most stable or *balanced* state of SO in the process of its evolution. In this case ($z_i = z_0$) — as it follows from (18) and (28),
$$\varepsilon = f\, M^{11/12}\,. \tag{34}$$
2. The number of m_g-elements is reduced to one, which will include all the m_i-structures. This state corresponds to the end of a certain period of object's evolution, i.e., to the *degenerate* state. Here, from (28),
$$\varepsilon = f M^{2/3}\,. \tag{35}$$

In the state of degeneration, when $z_0 = 1$, the period of core rotation will — as follows from (30), (32), (35) — be constant for any masses and amount to 2.51 sec, whereas the size of the outer layer will be equal to the standard radius $R_\odot$. In the general case, one can write, combining (34) and (35):

$$\varepsilon = fM^k, \tag{36}$$

where the parameter $k \geqslant 2/3$.

Visible dimensions of stars, i.e., radii of their photospheres, depend on many a specific factor; as a rule, they do not equal to the radius R_m and can be evaluated only roughly. The same can be said about star temperatures. Let us take the mass of the Sun as a standard (the validity of such a choice will be justified later) and consider the radius of the solar photosphere being close to $R_\odot$. Then, within the limits of the main sequence for the stable state and taking into account our disk model, the relative radius of the photosphere R_f for a star of arbitrary mass can be expressed via the mass of the Sun. It is evident that for a *two-dimensional* model,

$$R_f = \left(\frac{M}{M_\odot}\right)^{1/2} \tag{37}$$

and in the general case,

$$R_f = \left(\frac{M}{M_\odot}\right)^{i}, \tag{38}$$

where $i = 1 \ldots 1/3$ is a coefficient reflecting the density of packing of m_g-contours in the object.

To evaluate the model object temperature, let us consider its radiation as that of black body. Let the maximal temperature of radiation be achieved at the Compton wavelength of electron, $k = 2.426 \times 10^{-12}$ m, and let us assume that the radiation wavelength is inversely proportional to the rotary velocity of the contour vortex tubes at a given radius. Then, from Wien's formula,

$$T = \frac{b}{\lambda}, \tag{39}$$

where $b = 0.0029 \times 10^6$ m $\times\,^\circ$K. Having in mind this proportion, the radiation temperatures at the radii of core and photosphere (and an arbitrary radius as well) can be expressed as

$$T_0 = T_k\left(\frac{\mathrm{v}_0}{c}\right) \tag{40}$$

and

$$T_f = T_k\left(\frac{\mathrm{v}_0}{c}\right)\left(\frac{R_0}{R_f}\right), \tag{41}$$

whereas the energy of radiation (here and so forth, in keV) as

$$E = 511\frac{\mathrm{v}_0}{c}\ \text{keV}, \tag{42}$$

where T_k is the limiting temperature, corresponding to λ_k and equal to $1.19 \times 10^9\,^\circ$K.

Parameters	Balanced state	Degenerate state
ε	2.47×10^{37}	6.56×10^{38}
z	26.6	1
The core		
R_0	0.0126	0.0126
v_0	4.7×10^{-4}	0.0126
τ_0, sec	66.9	2.51
$T_0\,^\circ$K	5.6×10^5	1.5×10^7
The outer layer		
R_m	26.6	1
v_m	2.21×10^{-7}	1.57×10^{-4}
τ_m, sec	$3 \times 10^8 = 9.6$ years	$1.58 \times 10^4 = 4.4$ hours
$T_m\,^\circ$K	263	1.89×10^5
The photosphere		
R_f	1	1
$T_f\,^\circ$K	7050	1.89×10^5

Table 1: Note — radii and velocities are expressed as fractions of $R_\odot$ and c.

4 Model adequacy

It seems improbable that such a schematic and simple model would yield plausible results towards stellar objects. Yet it does. Let us calculate some parameters of a *solar-mass star*. The mass of the Sun equals to 2×10^{30} kg; in relative units, upon division by M_m, $M_\odot = 2 \times 10^{-6}$.

Table 1 shows the results of calculations according to the formulas given above.

In our notation, *angular momentum* of the Sun is equal to

$$0.4(2 \times 10^{30})\,\mathrm{v}_0\,R_0 = 0.4\,M_\odot^{23/12} M_m c\,R_\odot = \\ = 1.09 \times 10^{42}\ \text{kg m}^2/\text{sec}, \tag{43}$$

where the coefficient 0.4 takes account of the spherical shape of the body.

Comparing the calculated equilibrium-state parameters of this averaged standard object (a solar-type star) with the actual parameters of the Sun, one can see a close correspondence between their sizes, surface and core temperatures and periods of the solar cycle activity. The Sun's angular momentum is calculated with almost *perfect precision*.

By the end of evolution, upon reaching the degenerate state (at $z_0 = 1$), the periods of the inner (τ_0) and outer (τ_m) cycles diminish to their limits (Table 1). In this case, the single-thread spiral structure would flatten into a disk — thick as the size of the core (R_0) and radiating to the sector of the disk plane. The period of radiation will be $\tau_m = 4.4$ h; impulse duration, $\tau_0 = 2.5$ sec; and temperatures of the core and outer layer correspond to energies, 6.4 and 0.08 keV respectively.

The presence of m_i-zones in the m_g-contour will bring uncertainty into the period of radiation, which will be inversely proportional to the number of m_i-zones. For an object of the solar mass, the uncertainty in the period of impulses will amount to $\tau_m/z_i = 4.4\,h/26.6 = 598$ sec.

These parameters are typical and correspond well to the *x*-ray sources, *barsters*. For example, they perfectly fit the parameters of the X-ray source 3U 1820-30 in the globular cluster NGC6624 [5] etc.

Of course, the model presented here reflects only some essential features of stellar object structure. A stellar object can consist of toroids (balance of magnetic and gravitational forces), whose current-conducting elements rotate above the closed longitudinal axis of the tor (balance of magnetic and inertial forces), whereas the toroids themselves are oriented in the plane of the rotating disk (balance of gravitational and inertial forces). Such a system should hardly be stable. The core would rotate faster than the periphery, and the m_g-contours would coil up, with their kinetic energy transforming into other forms (and then, probably, transforming back). Describing such a system as a multiturn plane-spiral mechanical pendulum might be nave, yet in any case, there should take place an *oscillatory process of the object's gravimagnitodynamical structure*. Indeed, the paired dark spots in the equatorial zone of the Sun seem to be the outlet of m_g-contours — undergoing magnetic reversal and changing their intensity and polarity with the period of 11 years. Their registered quantity (from several to a hundred) does not contradict the calculated mean $z_0 = 26.6$.

Now let us calculate the *density of the SO core*. In the atoms of stellar matter (hydrogen, for the most part), substance circulates, according to our model, within $p^+ - e^-$ — contours with the mass $\varepsilon_0\, r_0$, and circulation speed cannot be higher than that of light [1].

At the same time, the magnitude of the charge e_0 is constant at any quantum number and equals to the momentum of the contour mass $\varepsilon_0\, r_0\, \mathrm{v}_0$. At $\mathrm{v}_0 \to c$, $r_0 \to r_{0\,min}$, therefore

$$r_{0\,min} = \frac{e_0}{\varepsilon_0\, c} = 1.65 \times 10^{-12}\ \mathrm{m}\,. \qquad (44)$$

The density of maximally condensed hydrogen atoms will amount (for a spherical volume) to

$$\rho_{max} = \frac{3m_H}{4\pi r_{0\,min}^3} = 8.82 \times 10^7\ \mathrm{kg/m^3}\,, \qquad (45)$$

where m_H is the mass of a hydrogen atom.

Now let us represent the mean density of the core matter as a ratio of the core mass to its cubic radius. Having in mind the corresponding expressions, one can see that the density is invariable and depends only on the gravitational constant:

$$\rho_0 = \frac{MM_m}{R_0^3} = \frac{M_m}{R_\odot^3} = \frac{1}{2\pi\gamma \times [\mathrm{sec}]^2} = \\ = 2.38 \times 10^9\ \mathrm{kg/m^3}\,. \qquad (46)$$

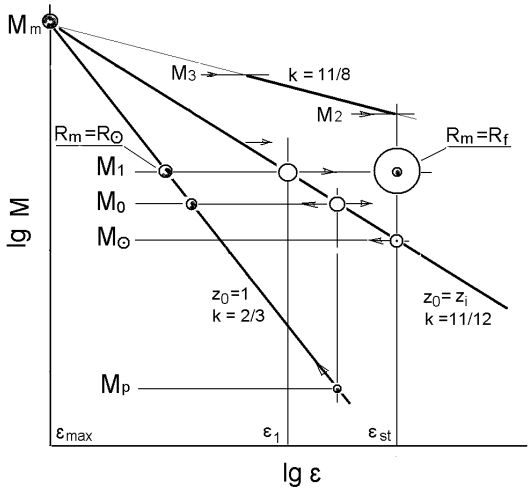

Fig. 1: The diagram "evolutionary parameter — mass".

As follows from the density ratio, a volume equal to that of a single hydrogen atom should contain 27 atoms of the initial matter, which corresponds, by the number of protons, to atoms of the iron group. The density is typical for white dwarfs, such as the famous Kuiper star.

It is interesting that the parameters obtained: $R_\odot$, ρ_0 and $\tau_0 = 2.51$ sec — practically indistinguishable from the values that should characterize the neck of a hypothetical magnetic "wormhole" of the mass M_m [4].

5 Analogues of the Hertzsprung-Russell (H-R) diagram and their applications

The Hertzsprung-Russell (H-R) diagram shows the evolutionary position of stellar objects on the "spectral class (temperature) — luminosity" coordinate plane. Let us consider its analogues: diagrams "evolutionary parameter — mass", and "temperature — mass".

5.1 The diagram "evolutionary parameter — mass"

On such a diagram (Fig. 1), $\varepsilon(M)$ dependencies would better be plotted on a logarithmic scale. At any k, the diagram rays converge on a point corresponding to the limiting mass M_m and limiting evolutionary parameter $\varepsilon_{max} = f$.

Specific parameters of SO will depend on the position of the object on the diagram. In general, with the converging point M_m approached then, as follows from (27–33), (40), (41), the number of m_g-contours will tend to 1; the rotary velocity, to the speed of light; the core and outer layer radii, to $R_\odot$; the periods of the inner and outer cycles, to 2.51 sec; and the core and outer layer temperatures, to T_k.

Evidently, for any given SO, the course of evolution may go both towards larger ε values (condensation of medium), up to $z = 1$, and smaller ε values (depression of medium), up to the shedding of the envelope at the end of the evolutionary process. ***Using the microcosm analogies, one can compare***

these states to the Bohr and ionized atoms respectively.

Let consider a stellar object which is in the main sequence and has a value of the evolutionary parameter corresponding to the line of equilibrium at $k = 11/12$. At $\varepsilon = const$, the equilibrium mass M_0 will correspond to a smaller mass M_p on the line of degeneration, for which $k = 2/3$ and $z_0 = 1$ (Fig. 1). In this case, one can obtain a mass ratio from (34) and (35):

$$M_p = M_0^{11/8} . \tag{47}$$

Since the mass of the Sun is considered standard, we shall take the evolutionary parameter value on the line of equilibrium for the solar mass ε_{st} as standard too.

5.2 Collapsing red giants

At the end of their evolution, stars become red giants and then shed their envelope (transfer to the state of the core), turning to white dwarfs, neutron stars or, in the case of the largest masses, "black holes".

Let us consider a star of chosen characteristic mass, for which every m_g-contour on the line of equilibrium has the mass of the Sun, i.e., satisfying the condition $M_0 = z_0 M_\odot$. Taking into account (28) and (34), we obtain $M_0 = M_\odot^{4/5} = 2.76 \times 10^{-5} = 13.8$ s.m. (masses of the Sun). Let us calculate the typical mass of a white dwarf forming from the core of such a star. Let us assume that on the line of star equilibrium, its core (and, therefore, the mass M_p as well) are on the line of degeneration (Fig. 1). Then, having in mind (47),

$$M_p = M_0^{11/8} = M_\odot^{11/10} = 5.38 \times 10^{-7} , \tag{48}$$

which corresponds to 0.27 s.m.

After the envelope and core are separated, they can be considered discretely. Let the envelope evolve to a standard parameter ε_{st}, and the core delay at the critical stage of the transformation process. Combining these states, let take the white dwarf mass M_p be proportional to the number of m_g-contours z_p — of the total number of m_g-contours z_0 of the mass M_0 at ε_{st}:

$$M_p = \frac{M_0 z_p}{z_0} . \tag{49}$$

Having in mind (28), (34) and (48), one can find the number of m_g-contours in the core:

$$z_p = \frac{f M_0^{25/24}}{\varepsilon_{st}} = M_\odot^{-1/12} = 2.98. \tag{50}$$

Therefore, the total mass of the star will be equivalent to $M_0/M_p = M_0^{-3/8} = M_\odot^{-3/10} = 51.2$ white dwarf masses, which corresponds to the number of nucleons in the nucleus of iron (more precisely, if $z_p = 3$, then $M_\odot = 1.9 \times 10^{-6}$ and the number of "nucleons" is equal to 52). Here we see another analogy with the microcosm: ***a standard red giant, containing 52 white dwarf masses, and a white dwarf, containing three***

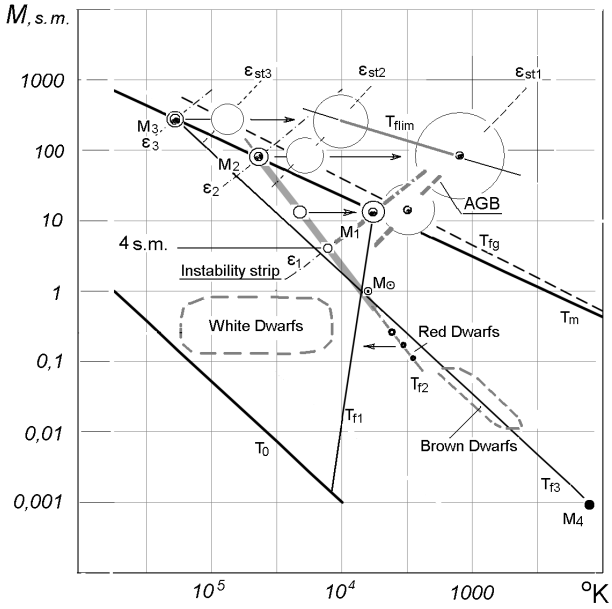

Fig. 2: The diagram "temperature-mass".

m_g-contours, will match an atom of iron, containing 52 nucleons, and a nucleon, consisting of three quarks. Later, other analogies with the microcosm will come into view.

Thus, it seems that the mass of the Sun and its evolutional parameter ε_{st} on the line of equilibrium are, indeed, standard. At $z_0 = 3$, the parameter $k \approx 0.75$, and it changes slightly in a wide range of masses. One can, therefore, expect that the condition (50) is optimal for other masses as well. Then, from (50),

$$\varepsilon_{st} = \frac{f M_0^{25/24}}{3} . \tag{51}$$

5.3 The diagram "temperature-mass"

Since logarithms of luminosity and mass are approximately proportional within the limits of the main sequence, it would be convenient to draw the H-R diagram analog in the coordinates of "temperature — mass".

From (27–30), (34), (40) and (41), one can obtain expressions of the $T(M)$ form, corresponding to the equilibrium temperatures at the radii of the outer layer R_m and core R_0 at $k = 11/12$. On a logarithmic scale (Fig. 2), they are straight lines, converging on the point M_m (outside the diagram):

$$T_m = T_k M^{7/6} , \tag{52}$$

$$T_0 = T_k M^{7/12} . \tag{53}$$

Stars of the main sequence have photospheres whose radii are usually smaller than R_m. To construct dependencies $T(M)$ for the photosphere, let us use formula (37). Taking into account (38), one can obtain, in the general case:

$$T_f = T_k M_\odot^i M^{k-i} . \tag{54}$$

For the equilibrium state at $k = 11/12$, we obtain three lines corresponding to three possible variants of packing of m_g-contours: into one-, two- and three-dimensional structures — i.e., at $i = 1, 1/2, 1/3$ (Fig. 2):

$$T_{f_1} = T_k\, M_\odot\, M^{-1/12}\,, \tag{55}$$

$$T_{f_2} = T_k\, M_\odot^{1/2} M^{5/12}\,, \tag{56}$$

$$T_{f_3} = T_k\, M_\odot^{1/3} M^{7/12}\,. \tag{57}$$

These lines converge on the point with coordinates close to the real parameters of the Sun, and their crossing with the outer-layer equilibrium line gives three characteristic masses: M_1, M_2 and M_3. The mass $M_1 = M_\odot^{4/5} = 13.8$ s.m., i.e., this mass also satisfies the condition $M_1 = z_0\, M_\odot$ and is equal to the mass of a red giant, which was calculated in the previous section. The mass $M_2 = 79.4$ s.m. is the largest possible mass for a main-sequence star. According to (47), this mass can give rise to an object whose mass will be 3 s.m., which corresponds to the maximal mass of a neutron star. The mass $M_3 = 277$ s.m. is the largest possible mass for a star with the most condense packing. According to our model, the structure of SO is two-dimensional; hence, stars of the main sequence are on the line T_{f_2} (bold line). Here, on the diagram *T-M*, one can also see isolines of the parameter ε, which, following (27–30), (41) and combining the constants, will look as

$$T_f = 6.86 \times 10^{-77}\, \frac{\varepsilon^2}{M^{2/3}}\,. \tag{58}$$

It should be noted that specific sequences of the globular-cluster stars formed from a medium with the same evolutionary parameter are also located along their own ε isolines.

When stars leave the main sequence and evolve towards lesser ε and T (to the right on the diagram), SO parameters change; particularly increasing is the envelope radius. Let us assume that beyond the line of equilibrium, $R_f = R_m$ (actually, the visible sizes of a star depend on many specific factors but we shall abstract from them in our model).

When calculating temperatures of the star envelopes (41), we implied that a part of the core radiation energy is transformed into other forms or spent in the star inner processes. But for the envelopes of giant stars, which are located to the right of the equilibrium line on the *T-M* diagram, formula (41) gives underrated results. The average density of giant stars is extremely low, and the energy of hot core radiation will insignificantly be absorbed by the rarefied atmosphere of these stars. In this case, to determine temperature of the photosphere, one can use the well-known formula for thermal radiation power, considering core as a radiation source:

$$N = \sigma T^4 S\,, \tag{59}$$

where σ is the Stefan-Boltzmann constant equal to $5.67{\times}10^{-8}$ W m^{-2}(°K)$^{-4}$. Having in mind the evident dependence of temperature on the linear size, the temperature of the photosphere can be expressed via the temperature of the core:

$$T_f = T_0 \left(\frac{R_0}{R_f}\right)^{1/2}. \tag{60}$$

Taking into account (27–30), (40) and accepting $R_f = R_m$, one can obtain, by analogy to (58),

$$T_f = 1.4 \times 10^{-55}\, \frac{\varepsilon^{3/2}}{M^{1/2}}\,. \tag{61}$$

This formula should be used when the star evolves beyond the equilibrium line and the radius of its envelope greatly increases. It is evident that the formula gives a bit overrated values of T_f. In Fig. 2, isolines plotted according to (61) are indicated as ε_{st}.

Taking into account (51) and substituting the ε_{st} expression in (61), one can obtain the line $T(M)$, along which stars turning into red giants are lined up:

$$T_{fg} = 0.192\, T_k\, M_0^{17/16}\,. \tag{62}$$

The parameters of stars with the masses M_1 and M_2 calculated for different ε values are shown in Table 2.

As for the "superstar" object, with the calculated mass $M_3 = 277$ s.m., its existence has been verified. The recently discovered star R136a1 has the following parameters: $M_0 = 265$ s.m., $R_f = 63R_\odot$ and $T_f \geqslant 40000$°K [7]. The calculated parameters of such a star — assuming it to be on the extension of the main sequence — are as follows: T_{f_2}, according to (56), is equal to 72500°K; ε from (61) is equal to 4.8×10^{38}; $R_f = R_m$ and, according to (30), is equal to $57R_\odot$. In other words, the object should be somewhere to the right of the main sequence line.

Located in the bottom part of the diagram are red dwarfs. Their typical parameters are the following: mass, $0.1, \ldots, 0.8$ s.m.; radius, $0.1 \ldots 0.85R_\odot$; temperature, below 3800°K [8, 9]. Since their radii are approximately proportional to their masses, they are on the line T_{f_1}, but their temperatures are lower, so it looks like they are on the extension of the main sequence. It is supposed that they evolve towards more condensed states, i.e., towards higher ε and T.

Lying on the lower segment of the T_{f_3} line are brown dwarfs. Their typical parameters are: mass, $0.012 \ldots 0.08$ s.m.; temperature, $3000 \ldots 300$ °K. Their radii change insignificantly over the range of masses and are approximately equal to that of Jupiter [10, 11].

At the very bottom of the diagram is the mass $M_4 = 1.95 \times 10^{-9}$ — the giant planet Jupiter. The temperature of its outer layer on the line T_{f_3} is equal, according to (57), to 123°K, i.e., it is close to the temperature of the outer atmosphere layers. The densities of Jupiter, brown dwarfs and the Sun are approximately equal; all these objects are near the line T_{f_3}.

Thus, all the types of SO are arranged logically on the *T-M* diagram.

Parameters	M_1 = 13.8 s.m		M_2 = 79.4 s.m		
	ε_1	ε_{st_1}	ε_2	ε_{st_2}	ε_{st_1}
ε	2.76×10^{38}	2.47×10^{37}	1.37×10^{39}	1.53×10^{38}	2.47×10^{37}
v_0	0.00219	0.000197	0.0061	0.00068	0.00011
R_0	0.0302	0.0302	0.0542	0.0542	0.0542
R_m	13.8	153.4	8.9	80	495
R_f	3.7	153.4	8.9	80	495
τ_0, sec	34.5	388	22.3	200	1242
τ_m, days	83	1.15×10^5	7	5037	1.2×10^6
τ_{mz}, days	6	752	0.78	63	2409
T_0, °K	2.6×10^6	2.34×10^5	7.2×10^6	8.07×10^5	1.3×10^5
T_m, °K	5710	3290	44000	21000	1370
T_f, °K	21200	3290	44000	21000	1370

Table 2: Note — radii and velocities are expressed as fractions of $R_\odot$ and c.

5.4 Variability of stellar objects

The types of variability of SO radiation are very diverse, and variability is intrinsic, to some degree, to all SO including the Sun. The most common type of variability is optical alternating variability (pulsations). According to our model, such pulsations are a natural result of the existence of oscillatory processes in the complex SO structure.

The most stable, in terms of amplitude and period of brilliancy oscillations, are pulsating stars of high luminosity — *Cepheids*, yellow giant stars [12, 13]. On the diagram *T-M*, their position would correspond to the mass M_1 on the equilibrium line T_m, where $R_f = R_m$.

Leaving the main sequence, stars become variable upon crossing the isoline ε_1 (*instability strip*), corresponding to the equilibrium parameter ε for the characteristic mass M_1. As follows from the diagram $T-\varepsilon$, the parameter ε decreases for masses larger than M_1 and increases for masses smaller than M_1 — until it reaches the isoline ε_1.

The masses of Cepheids are in the range 4 . . . 20 s.m. The minimal Cepheids mass is defined by the intersection of the isoline ε_1 and the line T_{f_2}, giving $M = 4.1$ s.m. which agrees with the value indicated in [14]. One should bear in mind that this intersection *point* on the diagram *T-M* corresponds to a *segment* on the diagram *ε-M* — from the line of equilibrium to ε_1. This segment corresponds to the initial period when the star begins to descend the main sequence. During this process, $R_f \to R_m$, which results in the star luminosity to grow. The growth is not reflected on the *T-M* diagram; on the diagram $H-R$, it corresponds to the initial segment of the star's evolutionary track.

Going on, stars evolve in the direction of lower ε values and reach the isoline ε_{st_1} (*asymptotic branch of giants, ABG*). The isoline corresponds to the equilibrium parameter ε_{st} for the standard solar mass (Fig. 1), under which the sizes of the star envelopes and the periods of their outer cycles reach their maxima. Located on ABG are *long-period variable stars* (with the period of brilliancy oscillations up to 1000 days), *semi-regular variable stars* (with the period of brilliancy oscillations up to 2000 days) and so on. Within the framework of our model, their variability can be explained not only by the existence of the outer layer period, τ_m, but also by a heterogeneity of their outer layer radiance [15, 16]. The heterogeneity results from the passage — along the star disk perimeter with the intervals of τ_{mz} — of hot (cold) zones, containing m_g-contours.

The calculated parameters R_m, T_m and τ_{mz} for M_1 (Table 2) are in a reasonable agreement with the averaged observation data for Cepheids at ε_1 and for long-period variables at ε_{st_1} [12, 17].

The parameters of SO of the mass M_2 on the line of equilibrium at ε_2 approximately correspond to those of hot supergiants PV Tel-type, with the period of pulsations from 0.1 to 1 day. On the line T_{fg} at ε_{st_2}, they correspond to the parameters of α Cyg-type super-giants, with the periods from several days to several weeks [12]. Further evolution of such stars in the direction of smaller ε values results in the formation of red super-giants.

6 Compact stellar objects

This group of SO includes white dwarfs, having the maximally compact packing of atoms, with the density ρ_0, and stellar bodies based on neutron stars, whose matter is compressed to the nuclear density ρ_j. Such objects are formed in the extreme cases, when SO evolve in the direction of either the largest ε values (when $R_f \to R_\odot$; "outer-layer state") or the smallest ones (when the envelope is shed; "core state"). In both cases, the initial oscillatory process is replaced with the rotation of the final compact object, of the mass M_p, with

the rate v_p.

At the final stage of evolution, there is, as indicated in [18], the possibility of a physical "coupling" of the star envelope with the core. Let us assume that there exists a ***process analogous to the absorption of an electron by the proton; i.e., the final compact object acquires the momentum of the outer layer, with the transition to an "excited" state***. We cannot consider the mechanism of this phenomenon within the framework of our model (moreover, the envelope and the core are considered here as different states of the same single object), so let us restrict ourselves to a formal application of the momentum conservation law:

$$M_0 v_m = M_p v_p. \tag{63}$$

6.1 White dwarfs

A white dwarf resulting from the star evolution towards lesser ε values, should inherit the parameters of the star core by the moment of the envelope shedding. For a star of the mass M_1 the parameters will be as follows: core temperature, 234000°K; period of rotation, 388 sec (Table 2). According to (47), (27) and (46), the mass, radius and mean density of white dwarfs are 0.27 s.m., 0.0082 $R_\odot$ and 2.38×10^9 kg/m^3 respectively. Indeed, very young white dwarfs can be observed in the X-ray range; the periods of their pulsations are in the range of tens to thousands of seconds, and they have typical sizes and densities being in agreement with the calculated parameters [12, 19, 20].

A white dwarf resulting from the evolution of a low-mass star towards larger ε values (without shedding of the envelope) should have the mass $M_p \approx M_0$. Then, its $v_p \approx v_m$.

Having in mind (29), (31) and (36), let us represent v_m as

$$v_m = cM_0^{2k-2/3} \tag{64}$$

and the period of rotation as

$$\tau_m = \frac{R_0}{v_m} = 2.51 M_0^{1-2k}. \tag{65}$$

At $z = 1$ and $k = 2/3$, an object of the mass 0.27 s.m. will have the following parameters: $v_m/c = 6.7\times10^{-5}$; $\tau_m = 308$ sec; and the energy of radiation, according to (42), equal to 0.034 keV ($T = 79000$°K). Here, the calculated parameters are, too, typical for a young white dwarf. As the object on the *T-M* diagram shifts to the right, the parameter k grows, which corresponds to the decline of the rotary velocity and temperature of the white dwarf.

On the diagram "*spectrum-luminosity*", the zone of white dwarfs seems much narrower than that on the diagram *T-M*, since their luminosity is determined by the radius, which, according to (27), is proportional to cubic root of the object mass.

6.2 Neutronization

In the context of our model, the process of neutronization can be represented as a loss of stability of the structure of m_g-contours and the transition of the structure (through its inversion along the vertical axis) from the plain two-dimensional into a one-dimensional configuration, which is energetically more favorable. Let us assume that the result will be a single m_g-contour or just a single vortical tube (neutron object).

Roughly, the parameters of such a primitive object can be defined as in Chapter 3. Placing the parameter R along the vertical axis and considering $z = 1$, one can obtain:

$$v_n = \frac{fM_n c}{\varepsilon}, \tag{66}$$

$$d_n = \frac{\varepsilon^2 R_\odot}{f^2 M_n}, \tag{67}$$

$$R_n = \frac{fM_n R_\odot}{\varepsilon}, \tag{68}$$

$$\tau_n = 2.51\frac{(\varepsilon/f)^3}{M_n^2}. \tag{69}$$

Rotary velocity cannot exceed the speed of light. Therefore, at $v_n \leqslant c$, $\varepsilon \geqslant fM_n$. Thus, for compact objects, the parameter k in (36) should be $\leqslant 1$ (in any event, as follows from the comparison of the calculated and actual data, k cannot be much larger than 1). Let us limit ourselves to defining parameters at $v_n = c$. Expressing ε from (66), one can obtain:

$$d_n = M_n R_\odot, \tag{70}$$

$$R_n = R_\odot, \tag{71}$$

$$\tau_n = 2.51 M_n. \tag{72}$$

It should be noted that a high-frequency modulation with τ_n up to 10^{-6} sec is present on the radiation diagrams of some neutron stars — pulsars [6].

As the evolutionary parameter grows, the sizes of a neutron object shrink along the axes, and on the line of degeneration, at $z = 1$, one can rewrite expressions (67–69), having in mind (35), in the following form:

$$d_n = R_n = M_n^{1/3} R_\odot, \tag{73}$$

$$\tau_n = \frac{R_\odot}{c} = 2.51 \text{ sec}. \tag{74}$$

Of course, this scheme is ideal. In reality, the objects based on neutron stars are in some intermediate state, and in the general case,

$$d_n = M_n^j R_\odot, \tag{75}$$

where $j = 1/3, \ldots, 1$ is a coefficient taking account of the object packing (shape).

It seems that the neutron state should be realized, to some extent, in the core of any star — and this can be proved. Let

represent the mass of a single vortex tube as that of a cylinder of the length R_n and radius d_n. Taking into account (70) and (71),

$$M_n\, M_m = \rho_n\, (M_n R_\odot)^2\, R_\odot, \tag{76}$$

where ρ_n is the vortex tube averaged density. Let us assume that ρ_n cannot exceed the nuclear density ρ_j, which shall be considered equal to $m_p/r_e^3 = 7.47 \times 10^{16}$ kg/m^3, where m_p is the mass of a proton. Then, as follows from (76),

$$M_{min} \geqslant \frac{M_m}{\rho_j R_\odot^3}\,, \tag{77}$$

which, upon substitution of values, gives $3.19{\times}10^{-8} M_m$. This mass corresponds to 0.016 solar masses or 17 Jupiter masses — exactly what the smallest cosmological mass, which is still considered a star, should be.

6.3 Masses of "black holes"

The diagrams ε-M and T-M show the boundary of a critical mode, where the rotary velocity of a vortex tube reaches that of light. On the diagram ε-M, the ray indicating the critical situation looks — taking into account that M_n is the mass of the compact object to be raised — as

$$\varepsilon = f M_n = f M^{11/8}\,. \tag{78}$$

On the diagram T-M, the same ray has — upon substitution of ε in (61) — the following form:

$$T_{f\,lim} = T_k\, M^{25/16}. \tag{79}$$

As follows from this construction, a ray segment is limited by the ordinates of the masses M_2 and M_3 and intersection with the isolines ε_{st_1} and ε_{st_2} — there are almost perfect ternary points of intersection. It is these masses that give rise to neutron objects with the masses, according to (47), $3, \ldots, 16$ s.m., which are the sources of hard X-ray radiation and *candidates for the star mass "black holes"* [18].

Indeed, for giant stars of a mass $M_2{-}M_3$, the critical mode begins before the moment they reach the asymptotic branch of giants (super-giants). With further decrease of the parameter ε, a star should release the excess of angular momentum — probably, by means of dropping the excess mass, which can be interpreted as shedding of the envelope with the formation of *supernova*. Next, the star core of a mass $M_n < \varepsilon/f$ transforms to an object which presently is classified as the "black hole" candidate. If neutronization of SO occurs far beyond the critical boundary (at low ε values), the mass of the emerging object will be very small. The latter might be one of the causes of the supernova remnants to contain few compact objects.

6.4 Radio pulsars

In our model, the simplest radio pulsar is a vortex tube which, by definition, is in the region Y ("boson"). The vortex tube is a macro-oscillator or radiator, with oscillations forming as longitudinal vibrations along the entire tube, while propagating to the X region as a cross wave from their source (the entrance of the vortex tube to the Y region; orifice) [2]. Presumably, radiation in the observable region X has a wavelength λ_p commensurable with the characteristic size of a single element of the vortex tube. A vortex tube, according to (22), consists of $n = \varepsilon$ single vortex threads — therefore, the characteristic linear size of a single element (region of radiation) will amount, under the condition of maximally compact packing of vortex threads in three dimensions, to

$$d_p = \varepsilon^{1/3} r_e\,. \tag{80}$$

The speed of vortex tube rotation can be expressed as a proportion of light speed — using the analogies described in Chapter 3:

$$\mathrm{v}_p = c\,\frac{\lambda_k}{d_p}\,. \tag{81}$$

Taking into account (36) and combining the constants, one can find the period of a pulsar:

$$\tau_p = \frac{d_p}{\mathrm{v}_p} = \frac{\varepsilon^{2/3} r_e^2}{c\lambda_k} = 282.5\, M_p^{2k/3}\ \mathrm{sec}\,. \tag{82}$$

Along the vortex tube of the pulsar, radiation is formed by m_i-zones, the number of which is determined by the pulsar mass. The averaged profile of the radiation pulse is a result of random superposition of many single pulses. Therefore the duration of the generalized pulsar pulse τ_{p_i} can be in the range from the duration of a single m_i-zone pulse to the total duration of pulses of all the zones, i.e. from r_i/v_{0i} to $z_i\, r_i/\mathrm{v}_{0i}$. Having in mind (18), (19) and (21),

$$\tau_{pi} = 2.51 M_p^{1/2\ldots1/4}\,. \tag{83}$$

For a pulsar, the standard mass is taken as 1.4 of that of the Sun. Then the pulsar period at $k = 2/3 \ldots 1$ will be, according to (82), in the range from 0.97 to 0.045 sec; and the duration of the generalized pulse will be, according to (83), in the range from 0.1 to 0.0042 sec, this corresponding to the temporal parameters of the majority of radio pulsars [21–23].

Radio radiation of pulsars covers a broad range and is extremely heterogeneous in time, intensity and frequency. Nevertheless, there are stable averaged spectra of energy distribution over frequency obtained by multiple instant measurements of radiation at different frequencies over large periods of time.

Let $\lambda_p = 2\pi d_p$, then the *frequency of radiation*, taking into account (80–82), will be as follows:

$$\nu_p = \frac{c}{2\pi d_p} = \frac{c}{2\pi\varepsilon^{1/3} r_e}\ \mathrm{Hz}, \tag{84}$$

which, having in mind (82), can be reduced to

$$\nu_p = 1.77\tau_p^{-1/2}\ \text{GHz}. \tag{85}$$

Since d_p is the minimal parameter provided that m_i-zone are packed most compactly, expression (84) will give *maximal* frequencies. However, the averaged spectrum extends far in the region of low frequencies and has an energy maximum. On the basis of our model, this fact can be accounted for by pulsation of the vortex tube in the limits of d_n, formula (75), and the existence of its optimal packing, less than 3, which the pulsar assumes most of the time. As indicated in [3], it may be the fractal dimension $e = 2.72$. In this case, the power of the parameter ε will be equal to $1/e$, and, as follows from (84), $\nu_p/\nu_m = \varepsilon^{0.0345}$. Having in mind (82), one can obtain, for the frequency of the maximum:

$$\nu_m = 0.0804\tau_p^{-0.55}\ \text{GHz}. \tag{86}$$

Formulas (85) and (86) are virtually identical to the interpolation formulas given in [23].

Although radiation of pulsars is not thermal, the *power of radiation* N_p can be determined on the basis of a formal use of the Boltzmann formula for thermal radiation of black body under the following conditions:

- taken as the area of the radiating surface is the cross-section of the vortex tube, $S = d_p^2$;
- taken as the effective temperature T_{ef} is the temperature corresponding to the radio frequency T_ν increased proportionally to the relative length of the vortex tube (i.e. proportionally to the ratio of the initial-object* radius to the diameter of the vortex tube, $T_{ef} = T_\nu R_0/d_p$).

Since, having in mind (39, 40), $T_\nu = T_k\lambda_k/d_p$, one can obtain, taking into account (36) and (80) and combining the constants,

$$T_{ef} = 1.06\times 10^7 M_p^{1/3-2k/3}. \tag{87}$$

Finally, after calculating the constants, we get an expression for N_p:

$$N_p = \sigma T_{ef}^4\, S = 1.45\times 10^{20} M_p^{4/3-2k}\ \text{W}. \tag{88}$$

Thus, our model predicts that at $k \to 2/3$, a radio pulsar should have a *lower* limit for radiation power (N_{min}), which the pulsar will be approaching as its rotation is getting slower. The limit N_{min} is equal to 1.45×10^{20} W and does not depend on the pulsar mass. At $k = 1$, expression (88) will give an upper limit N_p, which is dependent on the pulsar mass. The limits do exist [23], and no pulsars has been found at the luminosity below N_{min}.

On the basis of (82) and (88), a dependence $N(\tau_p)$ can be constructed (Fig. 3), which corresponds to the correlation given in [23]. To cover the zone of millisecond pulsars, the dependence is plotted in the range of masses $3\ldots 0.016$ s.m. — i.e. up to the minimal masses still able to neutronize (see Chapter 6.2). (The question on the range of radio pulsar masses is still open, since they can be determined only in rare cases).

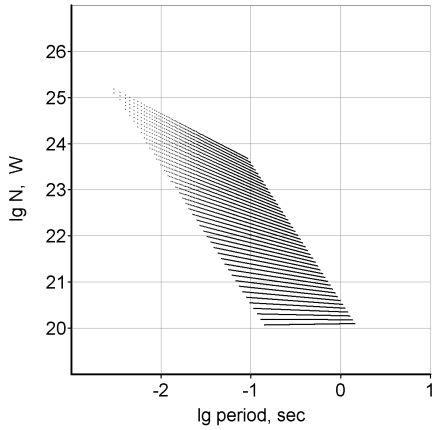

Fig. 3: Dependence of the radio pulsar radiation power on its period. $M_p = 3\ldots 0.016$ s.m., $k = 0.66\ldots 1$.

6.5 Excited states. Gamma-pulsar

Essentially, pulsar or vortex tube is a lengthy solenoid. In our model, the full length of a thread $z_i\, l_i$ does not depend, according to (18) and (20), from the mass and is equal to $R_\odot$; the length of a turn is, in general case, $\pi M_p^j R_\odot$, and the number of turns in the initial state is $\mathrm{N} = M^{-j}/\pi$.

Let us assume that the configuration of the vortex tube can change — e.g., upon the formation of a secondary spiral structure. In this case, the initial radius can diminish to the minimal radius of the vortex tube d_p, and the number of turns can grow to the number $\mathrm{N}_m = R_\odot/\pi d_p$. Then, taking into account (36) and (80),

$$\frac{\mathrm{N}_m}{\mathrm{N}} = 1.66\times 10^9 M_p^{j-k/3} = 10^5\ldots 10^9, \tag{89}$$

which will result in the correspondingly increased magnetic power and activity of the pulsar.

This state can be considered as an "excited" state of the radio pulsar. If the effective temperature grows proportionally as well, the energy corresponding to this increase will be transferred into the gamma range. Multiplying (87) by (89) and taking into account that for the vortex tube $j = 1$, one can obtain

$$T_{ef} = 1.76\times 10^{16} M_p^{4/3-k}. \tag{90}$$

Thus, at certain combinations of the parameters, formula (90) will give (upon conversion into electron-volts) values up to $10^{13}\ldots 10^{14}$ eV. This explains, for example, the observed gamma radiation of the famous pulsar in the Crab Nebula (more than 10^{12} eV). Ratio (89) serves estimation purposes, yet it can be used in other cases as well.

*The object of the initial mass (before neutronization).

6.6 X-ray pulsars

Massive stars give rise to neutron objects. Let us assume that such an object can be formed at any stage of star evolution, with the envelope momentum transferred to this newly formed object. Let us also assume that further evolution of this system as a whole can go both to the right (up to the state of outer layer) and left (up to the state of core) of the equilibrium line with the eventual formation of an *x*-ray pulsar of the mass M_p.

As a rule, X-ray pulsars do not radiate in the radio range. According to the model considered, we can assume this residual compact object to be already in the neutron state, while its vortex tube (or a part of the tube) excited at the expense of an additionally absorbed momentum to be still in the X region and to radiate in the X-ray range.

Let us determine the pulsar's parameters. Having in mind (63) and (64) and substituting, according to (47), $M_p^{8/11}$ for M_0, one can obtain for the pulsar:

$$\mathrm{v}_p = cM_p^{1.454k-0.7575}\,, \tag{91}$$

$$E_p = 511M_p^{1.454k-0.7575}\ \mathrm{keV}\,. \tag{92}$$

The pulsar period d_n/v_p, in the case of arbitrary pulsar form, will be equal to

$$\tau_p = \frac{M_p^j R_\odot}{\mathrm{v}_p} = 2.51\, M_p^{0.7575-1.454k+j}\,. \tag{93}$$

It should be noted that at $k = 0.75$ and $j = 1/3$, the M_p factor in (93) will be zero and $\tau_p = 2.51$ sec — the same period for any mass.

Let us consider the pulsar radiation to be mainly thermal. Then, one can calculate its power according to the Boltzmann formula, taking as the *radiating surface* that of the vortex tube of the length R_0 (i.e. $S = \pi d_p R_0$). In this case — analogously to (88), taking into account (27), having in mind $T_p = (T_k E_p)/511$ and after transformations — one can obtain, for an X-ray pulsar:

$$N_p = 1.22\times 10^{38} M_p^{6.15k-2.7}\ \mathrm{W}\,. \tag{94}$$

The parameters of most of the known X-ray pulsars fit into the intervals calculated according to (92–94) for the standard mass 1.4 s.m. at $k = 2/3\ldots1$ and $j = 1/3\ldots1$: $\tau_p = 0.002\ldots260$ sec, $E_p = 0.07\ldots35$ keV, $N_p = 10^{20}\ldots10^{30}$ W. Periods of more than 1000 sec are characteristic for small masses or for the cases when momentum is not fully transferred from the outer layer to the emerging compact object. Thus, there exist restrictions on the magnitudes of periods, energy and radiation power; and it is them that explain, to a certain degree, the partially non-thermal form of the pulsars energy spectrum (a cut-off in its high-energy region) [18, 24].

Radiating in the X-ray region are also some radio pulsars. Let us demonstrate the adequacy of our model on these objects — on the example of correlation between *x*-ray luminosity and the parameter $(d\tau/dt)/\tau^{3.5}$, given in [23]. The period derivative $d\tau/dt$, the rate of deceleration of pulsar rotation, is determined from observations. In our model, rotation slowdown is determined by the general process of evolution of the object's medium, i.e., by the parameter k. So let us use a derivative of the period in respect to k, considering the parameter j constant and replace the aforementioned expression by corresponding equivalent. In the end, differentiating (93) and combining the constants, one can obtain

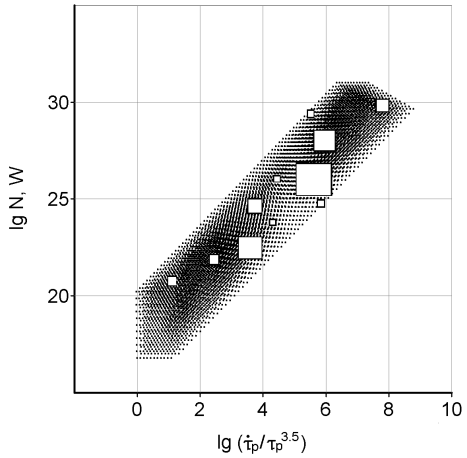

Fig. 4: Dependence of the radio pulsar *x*-ray luminosity on the parameter $(d\tau_p/dk)/\tau_p^{3.5}$. $M_p = 3\ldots0.3$ s.m., $k = 0.66\ldots1$, $j = 0.68\ldots0.73$. Observation data are taken from [23].

$$\frac{d\tau_p/dk}{\tau_p^{3.5}} = -3.35\lg M_p\,\tau_p^{-2.5}. \tag{95}$$

Fig. 4 shows the dependence of X-ray luminosity of a radio pulsar on the parameter $(d\tau_p/dk)/\tau_p^{3.5}$ in the range of masses $3\ldots0.3$ s.m. The dependence fits the observation data at the values of the parameter $j = 0.68\ldots0.73$. In Fig. 4, the size of squares is approximately proportional to the number of observation points (41 points in total according to [23]). In our case, the derivative does not require a scale coefficient to satisfy the initial conditions.

It is known that during *outbursts*, the power of radiation (luminosity) reaches a magnitude of the order of 10^{32} W and higher [25]. According to our model, such an increase in luminosity can be explained by periodical excitation of the vortex tube (see Section 6.5). In this case, multiplying (94) by (89), one can obtain

$$N_{pm} = 2.03\times 10^{47} M_p^{5.82k+j-2.7}\,. \tag{96}$$

Formula (96) gives rational results. For the mass $M = 1.4$ s.m., N_p will reach, depending on the parameters, magnitudes of $10^{38}\ldots10^{39}$ W, which agrees with the power of the giant gamma-ray outburst from the source SGR 1900-14, which was registered in August 1998 (about 10^{38} W) [27].

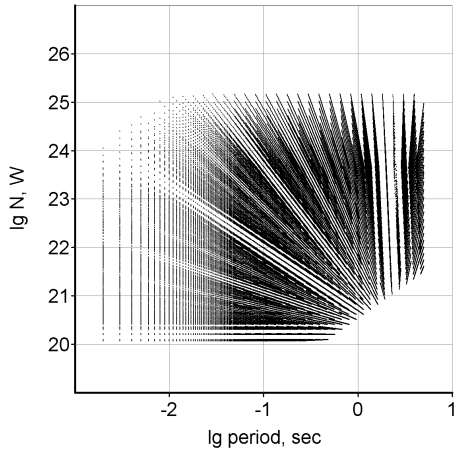

Fig. 5: The solution region: dependence of the radio pulsar radiation power on its period. $M_p = 3 \ldots 0.016$ s.m., $k = 0.66 \ldots 1$, $j = 0.33 \ldots 1$.

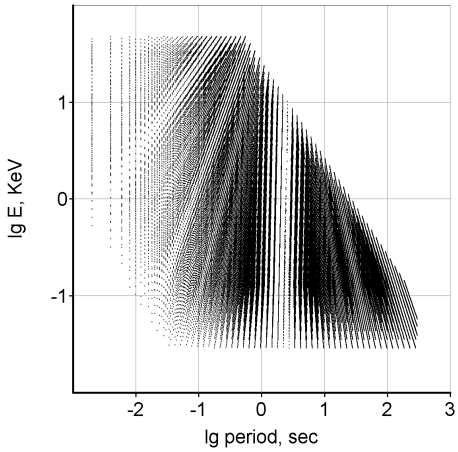

Fig. 6: The solution region: dependence of the X-ray pulsar radiation energy on its period. $M_p = 3 \ldots 0.3$ s.m., $k = 0.66 \ldots 1$, $j = 0.33 \ldots 1$.

It would be interesting to get independent estimates of the mass of compact objects, which, as one can see, have a similar origin. Let us assume that in the process of their possible inter-transformations, their masses and periods change insignificantly. Let the X-ray and radio pulsar periods are equal in the marginal cases — when the initial SO, giving rise to a compact object, evolves towards the largest or smallest ε values.

Let us consider the case when evolution goes towards larger ε. With ε increasing, the mass M_p should grow and at $z \to 1$ become equal to the original mass M_0 (Fig. 1). Perhaps, such a process should be associated with *accretion in binary star systems*. Proceeding to the mass M_0, let us substitute $M_0^{11/8}$ for M_p in (91). Then $d_n = M_0^j R_\odot$ and (93) will take a form of

$$\tau_p = 2.51 M_0^{1.042-2k+j}. \tag{97}$$

Equating (82) to (97) for the periods, combining the constants and making transformations, one can obtain in the end:

$$\lg M_0 = \frac{2.052}{1.042 - 2.667k + j}. \tag{98}$$

In the limit, $k = 2/3$ and $j = 1/3$ (sphere), then $M_0 = 8{\times}10^{-6}$ or 4 s.m. This mass can be considered as the total one of a *low-mass binary star system* containing an X-ray pulsar, this being in agreement with the accepted estimate (2.5 s.m. + 1.4 s.m.) [18]. Such a pulsar will have a relatively hard X-ray radiation [25], and, with the growth of the parameters j, its period will decrease.

The obtained mass value is, in fact, coincides with the minimal mass of a Cepheids (see Section 5.4). Thus, an SO with the mass 4 s.m. can evolve both to the right of the equilibrium line (shedding the envelope) and to the left (forming a binary star system). In both cases, a compact object will be formed at the end of evolution, and one can suppose that the stellar mass of 4 s.m. is the *minimal mass* able to give rise to neutron stars.

Let an X-ray pulsar evolve towards lesser ε values. Equating expressions (82) and (93), one can obtain

$$\lg M_p = \frac{2.052}{0.7575 - 2.121k + j}. \tag{99}$$

In the limit, $k = 1$ and $j = 1$ (vortex tube), then $M_p = 2.3 \times 10^{-6}$ or 1.15 s.m. Here, we have got a typical pulsar mass. Such a pulsar will have a relatively soft X-ray radiation, and with the parameter j growing, the pulsar period will increase. Such objects can correspond to *single neutron stars* [26]. Indeed, as follows from the observation data, pulsars of binary systems will mainly speed up their rotation, whereas single objects will slow down.

The properties of SO are determined by the totality of their parameters; that is why two-parameter diagrams always have a wide scatter of experimental points. Let us represent the solution region of the dependence $N(\tau_p)$ for radio pulsars more extensively — expressing its period according to (93), which contains the parameter j, and considering some radio pulsars evolved from the X-ray ones, with their periods being approximately the same (Fig. 5). The region of observation values [23] fits well the solution region.

Analogously, using formulas (92) and (93), one can plot a solution region of the dependence $E(\tau_p)$ for the X-ray pulsars (Fig. 6). Clusters on the images may indicate regions where pulsars have preferable parameters — e.g., the right bottom part in Fig. 6 may indicate, by the combination of parameters, a region of single neutron stars.

There appears a question: can slow X-ray pulsars transform into radio pulsars, whose period will not exceed several seconds? One can suppose that comparatively to radio pulsars, X-ray ones have an excessive angular momentum (since their radius in the region X is much larger than that of ra-

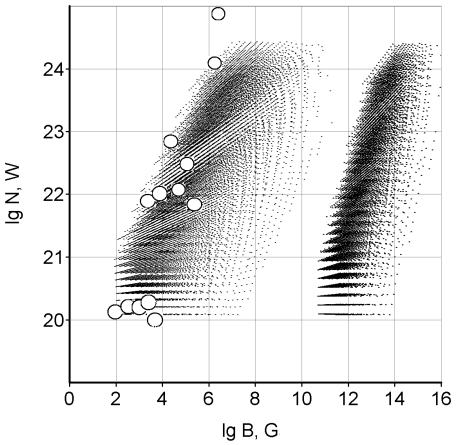

Fig. 7: The solution region: dependence of the radio pulsar radiation power on its magnetic field — $N(B)$ to the left; $N(B_m)$ to the right. $M_p = 2 \ldots 0.2$ s.m., $\tau_p = 0.003 \ldots 3$ sec, $k = 0.66 \ldots 1$, $j = 0.33 \ldots 1$. Observation data are taken from [23].

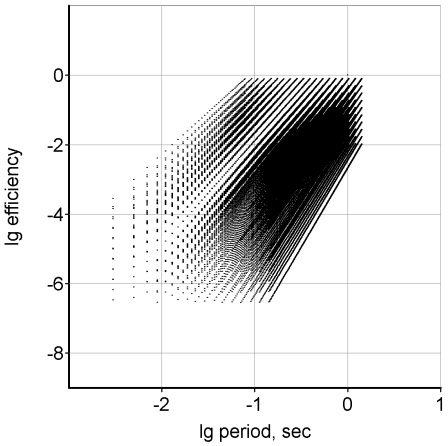

Fig. 8: The solution region: dependence of the efficiency of transformation of rotation energy in-to radio radiation on the pulsar period (initial state). $M_p = 3 \ldots 0.016$ s.m., $k = 0.66 \ldots 1$, $j = 0.33 \ldots 0.55$.

dio pulsars in the region Y, and as they "submerge" into the region Y, their period shortens).

Thus, ***it can be supposed that gamma, X-ray and radio pulsars are different forms of excited vortex tube or, using another analogy with the microcosm, three species of neutrino. The primary state — radio pulsar — possesses only the initial angular momentum of the vortex tube or spin***.

6.7 Magnetic properties of pulsars

Our model explains the correlation between the magnitude of the magnetic field B and other pulsar parameters. According to SI definition, for a lengthy solenoid, $B = \mu\mu_0\, nI$, where n is the number of turns per unit of length, I is the current strength and μ is the relative magnetic permeability.

The initial solenoid length is equal to R_0. Let $n = \mathrm{N}/R_0$. Let us define the coefficient μ as the compactness of the solenoid coil in the initial state $\mathrm{N}d_p/R_0$. The current strength I in the "Coulombless" form is $z_e m_e\, c(R_\odot/r_e) \times 1/[\mathrm{sec}]$ (see Section 2), where z_e is the number of single charges per coulomb, equal to $1/e_0$.

In our model, SI units for B are m^{-1}. To switch from SI to the Gaussian system of units, introduction of an additional factor of 10^{-4} is needed. Opening the expressions for μ_0, ε_0 and $R_\odot$, taking into account that $\mathrm{N} = M^{-j}/\pi$, as well as (27), (36) and (80), and making transformations, one can finally obtain

$$B = 1.27 \times 10^{-4} M_p^{k/3-2j-2/3}\ \mathrm{G}. \qquad (100)$$

Many radio pulsars have larger B values. For the excited state, multiplying (100) by (89), we will have

$$B_m = 2.1 \times 10^{5} M_p^{-j-2/3}\ \mathrm{G}. \qquad (101)$$

Fig. 7 shows the solution regions for the dependences $N(B)$ (to the left) and $N(B_m)$ (to the right) calculated according to formulas (88), (100) and (101) in the range of masses $2 \ldots 0.2$ s.m. and periods $0.003 \ldots 3$ sec. The figure also represents the observation data for the pulsars with small B values taken from [23]. Masses and periods are connected using formula (93), which contains the parameter j. It is known that according to the strength of their magnetic field, pulsars are clustered near values of the order of 10^9 and 10^{13} G [18], which agrees, in general, with the distributions obtained.

To analyze pulsar parameters, the function $\eta(\tau_p)$ is also used, which includes the magnetic force B [23]:

$$\eta = \frac{3N_p\, c^3 \tau_p^4}{8\pi^4 B^2 R_*^6}, \qquad (102)$$

where η is the pulsar efficiency, i.e., the effectiveness of transformation of the pulsar rotation energy into radio radiation.

According to [23], formula (102) takes $R_* = 10^6$ cm. For more objectiveness, let us replace this constant with the diameter of the vortex tube according to (75). Having in mind (82), (88) and (100), let us transform (102) to the form (in the Gaussian system):

$$\lg \eta = 8.5 + (2.667 - 2j) \lg M_p\,. \qquad (103)$$

Together with formula (82), this gives the region of $\eta(\tau)$ solutions for radio pulsars (Fig. 8). Since $\eta < 1$, there are limitations for some combinations of the parameters. In the accepted, according to [23], range of η values, the parameter j is limited by the range $0.33 \ldots 0.55$, which is characteristic for pulsars with small B values. The orientation of clusters on the diagram indicates the increase of η with the growth of the period.

Analogously, substituting the parameter B_m into (102), one can obtain

$$\lg \eta = -11.9 + (0.667k - 4j + 2.667) \lg M_p. \qquad (104)$$

In this case (Fig. 9), in the accepted range of η values, the parameter j is limited by a narrow range of large values,

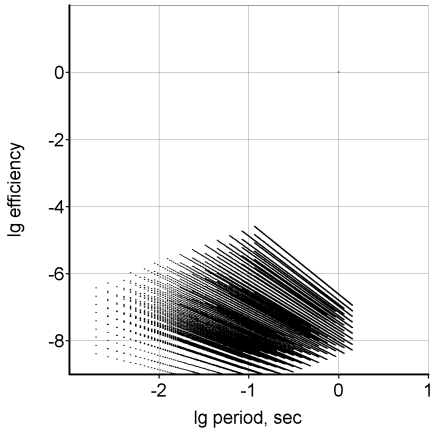

Fig. 9: The solution region: dependence of the efficiency of transformation of rotation energy in-to radio radiation on the pulsar period (excited state). $M_p = 3 \ldots 0.016$ s.m., $k = 0.66 \ldots 1$, $j = 0.94 \ldots 1$.

$0.94 \ldots 1$, which corresponds to pulsars with a strong magnetic field. In this range of parameters, η will grow as the period decreases. These solution ranges complement each other and agree with the body of the observation data of the diagram $\eta(\tau)$ given in [23]. Thus, there are at least two pulsar populations, with different magnitudes of their magnetic field and different form factors (parameter j), which was also indicated in [23].

From (101), one can find that the magnitude of the pulsar magnetic field can reach $10^{14} \ldots 10^{15}$ G. Such a growth of the magnetic field also explains the phenomenon of *magnetars* [27, 28].

As follows from our model — and it is getting evidence now — there are no essential differences between magnetars and X-ray pulsars. For example, the sources SWIFT J1822.31606 [29] and PSR J18460258 [30] possess features of both objects.

7 Conclusion

Thus, our model, which is built exclusively on the balances of basic interactions, describes different kinds of stellar objects. It is shown that SO features are mainly determined by their masses and the state of the evolving medium that they are made of. Together with the basic constants, these parameters (M and ε) determine the evolutionary behavior of stellar objects and the very existence of the well-known Hertzsprung-Russell diagram. In a number of cases, they are sufficient for the calculation of basic SO parameters: the mass of the final compact objects, radiation energy, radiation power and periods or rotation.

The model reveals analogies between the macro- and microlevels of matter: cosmological masses and elementary particles.

Indeed, the ***general range of stellar masses can be roughly divided into three subranges — by the analogy with the three families of elementary particles***:

- stars with masses less than 4 s.m., which in the end of evolution will become white dwarfs;
- giant stars with masses $4 \ldots 79$ s.m., which in the end of evolution will give raise to neutron stars;
- super-giant stars with masses $79 \ldots 277$ s.m., which in the end of evolution will give raise to X-ray sources — candidates for black holes.

It is the ***stars of small masses and their final states (cold white dwarfs, "protons") that are the "first family" of stellar population***. They make the majority of it and are stable on the cosmological scale, since their lifetimes are immeasurably longer than the lifetimes of other stellar objects.

Hopefully, the results obtained and the presented model can be useful for further theoretical studies in the field.

Submitted on: June 26, 2012 / Accepted on: October 10, 2012

References

1. Belyakov A.V. Charge of the electron, and the constants of radiation according to J. A. Wheeler's geometrodynamic model. *Progress in Physics*, 2010, v.4, 90–94.
2. Belyakov A.V. Macro-analogies and gravitation in the micro-world: further elaboration of Wheeler's model of geometrodynamics. *Progress in Physics*, 2012, v.2, 47–57.
3. Belyakov A.V. On the independent determination of the ultimate density of physical vacuum. *Progress in Physics*, 2011, v.2, 27–29.
4. Novikov I.D., Kardashev N.S., Shatskiy A.A. Multicomponent Universe and astrophysic of the "wormhole". *Uspekhi-Physics*, 2007, v.177(9), 1017–1023.
5. Beck P.G., Montalban J. et al. Fast core rotation in red-giant stars as revealed by gravity-dominated mixed modes. *Nature*, January 2012, v.481, 55–57.
6. Narlikar J. Violent Phenomena in the Universe. Oxford University Press, Oxford, 1984.
7. Crowther P.A. The R136 star cluster hosts several stars whose individual masses greatly exceed the accepted 150 $M_\odot$ stellar mass limit. *Monthly Notices of the Royal Astronomical Society*, 2010, v.408(2), 731–751.
8. Burrows A., Hubbard W.B., Saumon D., Lunine, J.I. An expanded set of brown dwarf and very low mass star models. *The Astrophysical Journal*, 1993, v.406, no.1, 158–171.
9. Kaltenegger L., Wesley A. Transits of earth-like planets. *The Astrophysical Journal*, 10 June 2009, v.698, 519–527.
10. Spiegel D.S., Burrows A., Milsom J.A. The Deuterium-burning mass limit for brown dwarfs and giant planets. ariv: astro-ph/1008.5150.
11. Michael C.L. et al. CFBDSIR J1458+1013B: a very cold (>T10) brown dwarf in a binary system. ariv: astro-ph/1103.0014.
12. Kholopov P.N., Samus N.N., Goranskiy V.P. et al. Main catalog of the variable stars. v. I-III. Nauka, Moscow, 1985–1987 (in Russian).
13. Berdnikov L.N., Samus N.N. Studies of classical cepheids. *Astronomical & Astrophysical Transactions*, 1999, v.18(2), 373–384.
14. Berdnikov L.N., Dambis A.K. Cepheids and RR Lyr variables. Present star astronomy. Sternberg Astron. Inst., Moscow, 16 June 2011 (in Russian).
15. Mennessier M.O., Mowlavi N., Alvarez R., Luri X. Long period variable stars: galactic populations and infrared luminosity calibrations. *Astronomy & Astrophysics*, 2001, v.374, 968–979.

16. Hampton M., Gregory W.H. et al. HD 12545, a study in spottedness. *Publication of the Astronomical Society of the Pacific*, January 1996, v.108, 68–72.

17. Le Bertre T., Lebre A., Waelkens C. (Eds.). Asymptotic Giant Branch Stars. *Proc. 191st Symp. IAU*, 1999.

18. Svertilov S.I. Cosmic X-ray and gamma-ray radiation. Skobeltzin Inst. Nucl. Research, Moscow, April 2006 (in Russian).

19. Shapiro S.L., Teukolcky S.A. Black holes, white dwarfs and neutron stars. v.1–2. Nauka, Moscow, 1985.

20. Kepler S.O., Vauclair G., Nather R.E., Winget D.E., Robinson E.L. G117-B15A — how is it evolving? *White dwarfs; Proceedings of IAU Colloquium 114th*, August 1988, (A90-32719 13-90).

21. Manchester R., Talor D. Pulsars. Nauka, Moscow, 1980.

22. Matveenko L.I., Usov V.V. Physics of the Cosmos. Nauka, Moscow, 1986 (in Russian).

23. Malov I.F. Radio pulsars. Nauka, Moscow, 2004 (in Russian).

24. Tsygankov S.S., Lutovinov A.A. Studies of x-ray pulsars with space observatories. *Proceedings of the VIII-th Young Scientist Conference*, 2005, 226–228 (in Russian).

25. Aleksandrovich N.A., Borozdin K.N., Arefjev V.A., Sunyaev R.A., Skinner D.K. Observations of the X-ray transit pulsar-burster GRO J1744-28 with the telescope TTM of the orbital observatory "MIR-KVANT". *Soviet Astron. Journal Letters*, 1998, v.24, no.1.

26. Popov S.B., Prohorov M.E. Astrophysics of the isolated neutron stars: radio-quiet neutron stars and magnetars. Sternberg Astron. Inst., Moscow, 2002 (in Russian).

27. Kouveliotou C., Duncan R.C., Thompson C. Magnetars. *Scientific American Russian Edition*, 2003, no.6 (in Russian).

28. Tsygankov S.S., Lutovinov A.A. Universe magnetic heart. *Priroda*, 2011, no.1, 10–18 (in Russian).

29. Scholz P., Ng C.-Y., Livingstone M.A., Kaspi V.M., Cumming A., Archibald R. SWIFT J1822.31606: post-outburst evolution of a nearby magnetar. arXiv: astro-ph/1204.1034.

30. Gavriil F.P., Gonzalez M.E., Gotthelf E.V., Kaspi V.M., Livingstone M.A., Woods P.M. Magnetar-like emission from the young pulsar in Kes 75. *Science*, 28 March 2008, v.319, no.5871, 1802–1805.

Sampling the Hydrogen Atom

Norman Graves

192 Old Woosehill Lane, Wokingham, UK. E-mail: normangraves@btinternet.com

A model is proposed for the hydrogen atom in which the electron is an objectively real particle orbiting at very near to light speed. The model is based on the postulate that certain velocity terms associated with orbiting bodies can be considered as being affected by relativity. This leads to a model for the atom in which the stable electron orbits are associated with orbital velocities where Gamma is n/α, leading to the idea that it is Gamma that is quantized and not angular momentum as in the Bohr and other models. The model provides a mechanism which leads to quantization of energy levels within the atom and also provides a simple mechanical explanation for the Fine Structure Constant. The mechanism is closely associated with the Sampling theorem and the related phenomenon of aliasing developed in the mid-20th century by engineers at Bell labs.

Since the emergence of quantum theory just over a century ago every model that has been developed for the hydrogen atom incorporates the same basic assumption. From Niels Bohr through de Broglie and Schrödinger up to and including the Standard Model all such theories are based on an assumption first put forward by John Nicholson.

Nicholson recognised that the units of Planck's constant are the same as those of angular momentum and so he reasoned that perhaps Planck's constant was a measure of the angular momentum of the orbiting electron. But Nicholson went one step further and argued that Planck's constant was the fundamental unit or quantum of angular momentum and therefore the angular momentum of the orbiting electron could only take on values which were an integer multiple of Planck's constant. This allowed Bohr to develop a model in which the energy levels of the hydrogen atom matched those of the empirically developed Rydberg formula [1]. When the Bohr model was superseded Nicholson's assumption was simply carried forward unchallenged into these later models.

Nicholson's assumption however lacks any mathematical rigour. It simply takes one variable, angular momentum, and asserts that if we allow it to have this characteristic quantization then we get energy levels which appear to be correct. In so doing it fails to provide any sort of explanation as to just why such a quantization should take place.

In the mid-20th century a branch of mathematics emerged which straddles the boundary between continuous functions and discrete solutions. It was developed by engineers at Bell Labs to address problems of capacity in the telephone network. While at first site there appears to be little to connect problems of network capacity with electrons orbiting atomic nuclei it is the application of these mathematical ideas which holds the key to explaining quantization inside the atom.

In the 1930's and 40's telecommunications engineers were concerned to increase the capacity of the telephone network. One of the ideas that surfaced was called Time Division Multiplexing. In this each of a number of incoming telephone lines is sampled by means of a switch, the resulting samples are sent over a trunk line and are decoded by a similar switch at the receiving end before being sent on their way. This allowed the trunk line to carry more telephone traffic without the expense of increasing the number of cables or individual lines. The question facing the engineers at the time was to determine the minimum frequency at which the incoming lines needed to be sampled in order that the telephone signal can be correctly reconstructed at the receiving end.

The solution to this problem was arrived at independently by a number of investigators, but is now largely credited to two engineers. The so called Nyquist-Shannon sampling theorem is named after Harry Nyquist [2] and Claude Shannon [3] who were both working at Bell Labs at the time. The theorem states that in order to reproduce a signal with no loss of information, then the sampling frequency must be at least twice the highest frequency of interest in the signal itself. The theorem forms the basis of modern information theory and its range of applications extends well beyond transmission of analogue telephone calls, it underpins much of the digital revolution that has taken place in recent years.

What concerned Shannon and Nyquist was to sample a signal and then to be able to reproduce that signal at some remote location without any distortion, but a corollary to their work is to ask what happens if the frequency of interest extends beyond this Shannon limit? In this condition, sometimes called under sampling, there are frequency components in the sampled signal that extend beyond the Shannon limit and maybe even beyond the sampling frequency itself.

A simple example can be used to illustrate the phenomenon. Suppose there is a cannon on top of a hill, some distance away is an observer equipped with a stopwatch. The job of the observer is to calculate the distance from his current location to the cannon. Sound travels in air at roughly 340 m/s. So it is simply a matter of the observer looking for the flash as the cannon fires and timing the interval until he

hears the bang. Multiplying the result by 340 will give the distance to the cannon in metres, let's call this distance D.

This is fine if the cannon just fires a single shot, but suppose the cannon is rigged to fire at regular intervals, say T seconds apart. For the sake of argument and to simplify things, let's make T equal to 1. If the observer knows he is less than 340 m from the cannon there is no problem. He just makes the measurement as before and calculates the distance D. If on the other hand he is free to move anywhere with no restriction placed on his distance to the cannon then there is a problem. There is no way that the observer knows which bang is associated with which flash, so he might be located at any one of a number of different discrete distances from the cannon. Not just any old distance will do however. The observer must be at a distance of D or $D + 340$ or $D + 680$ and so on, in general $D + 340n$. The distance calculated as a result of measuring the time interval between bang and flash is ambiguous. In fact there are an infinite number of discrete distances which could be the result of any particular measured value. This phenomenon is known as aliasing. The term comes about because each actual distance is an alias for the measured distance.

Restricting the observer to be within 340 m of the cannon is simply a way of imposing Shannon's sampling limit and by removing this restriction we open up the possibility of ambiguity in determining the position of the observer due to aliasing.

Let's turn the problem around a little. If instead of measuring the distance to the cannon the position of the observer is fixed. Once again to make things simpler, let's choose a distance of 340m. This time however we are able to adjust the rate of fire of the cannon until the observer hears the bang and sees the flash as occurring simultaneously. If the rate of fire is one shot per second then the time taken for the slower bang to reach the observer exactly matches the interval between shots and so the two events, the bang and the flash are seen as being synchronous. Notice that the bang relates, not to the current flash, but to the previous flash.

If the rate of fire is increased then at first, for a small increment, the bang and the flash are no longer in sync. They come back into sync however when the rate of fire is exactly two shots per second, and again when the rate is three shots per second. If we had a fast enough machine gun this sequence would extend to infinity for a rate of fire which is an integer number of shots per second. Notice that now the bang no longer relates to the previous flash, but to a previous flash. It is interesting to note also that if the rate of fire is reduced from once per second then the observer will never hear and see the bang and the flash in sync with one another and so once per second represents the minimum rate of fire which will lead to a synchronous bang and flash. In fact what we have here is a system that has as its solutions a base frequency and an infinite set of harmonic frequencies.

Suppose now that there is some mechanism which feeds back from the observer to the cannon to drive the rate of fire such that bang and flash are in sync, and suppose that this feedback mechanism is such as to always force the condition to apply to the nearest rate of fire which produces synchronisation.

We now have a system which can cause a variable, in this case the rate of fire of the gun, to take on a series of discrete values even though, in theory at least, the rate of fire can vary continuously. Equally important is that if the feedback mechanism is capable of syncing the system to the lowest such frequency then all the multiples of this frequency are also solutions, in other words if the base frequency is a solution then so are harmonics of the base frequency.

This idea that there are multiple discrete solutions which are harmonics of a base frequency is an interesting one since it couples the domains of the continuous and the discrete. Furthermore what the example of the cannon shows us is that any system which produces results which are a harmonic sequence must involve some sort of sampling process. This becomes clear if we consider the Fourier representation of a harmonic sequence. A harmonic sequence of the type described consists of a number of discrete frequencies, spreading up the spectrum and spaced equally in the frequency domain with each discrete frequency represented by a so called Dirac function. Taken together they form what is described as a Dirac comb, in this case in the frequency domain. The inverse Fourier transform of such a Dirac comb is itself another Dirac comb, only this time in the time domain, and a Dirac comb in the time domain is a sampling signal [4].

This link between a Dirac comb in the frequency domain and a corresponding Dirac comb in the time domain means that if ever we observe a set of harmonics in some natural process there must inevitably be some form of sampling process taking place in the time domain and vice versa.

One such example, in which this relationship has seemingly been overlooked, is found in the structure of the hydrogen atom.

By the beginning of the 20th century it was becoming evident that the universe was composed of elements which were not smooth and continuous but were somehow lumpy or granular in nature. Matter was made up of atoms, atoms themselves contained electrons and later it emerged that the atomic nucleus was itself composed of protons and neutrons.

Perhaps even more surprising was that atoms could only absorb or emit energy at certain discrete levels. These energy levels are characteristic of the atom species and form the basis of modern spectroscopy. The issue facing the scientists of the day was that this discrete behaviour is not associated with the discrete nature of the structure of the atom; that can easily be explained by asserting that any atom contains an integer number of constituent particles. Where energy levels are concerned, the quantization effects involve some sort of process that is taking place inside the atom.

The atom with the simplest structure is that of hydrogen,

comprising a single proton surrounded by an orbiting electron and work began to investigate its structure and to understand the mechanisms which gave it its characteristic properties.

The first such theoretical model was proposed by Niels Bohr [5]. Bohr used simple classical mechanics to balance the centrifugal force of the orbiting electron against the electrostatic force pulling it towards the atomic nucleus. He needed a second equation in order to solve for the radius and velocity of the orbiting electron and came upon the idea proposed by John Nicholson [6]. Nicholson reasoned that the units of Planck's constant matched those of angular momentum and so he proposed that the angular momentum of the orbiting electron could only take on values which were an integer multiple of was Planck's constant.

Bohr's equations worked, but they threw up a strange anomaly. In Bohr's model each energy level is represented by the orbiting electron having a specific orbit with its own particular orbital velocity and orbital radius. The really strange thing was that in order to fit with the conservation laws, transitions from one energy state to another had to take place instantly and in such a way that the electron moved from one orbit to another without ever occupying anywhere in between, a sort of discontinuity of position. This ability to jump instantaneously across space was quickly dubbed the Quantum Leap in the popular media, a phrase which still has resonance today.

Bohr reasoned that

$$l = mv_n r_n = n\hbar \tag{1}$$

$$\frac{Kq^2}{\hbar c} = \frac{mv_n^2}{r_n} \tag{2}$$

which means

$$v_n = \frac{Kq^2}{n\hbar} \tag{3}$$

$$r_n = \frac{n^2\hbar^2}{mKq^2}\,. \tag{4}$$

where m is the rest mass of the electron, q is the charge on the electron, r_n is the orbital radius for the nth energy level, v_n is the orbital velocity for the nth energy level, l is the angular momentum, K is the Coulomb force constant, $\hbar$ is Planck's constant.

Equation 1 represents Nicholson's assumption that angular momentum can only take on values which are integer multiples of Planck's constant.

Equation 2 balances the centrifugal force against the electrostatic force.

Equation 3 shows that the orbital velocity decreases with increasing energy level.

Equation 4 shows that the orbital radius increases as the square of the energy level and leads directly to the idea of the Quantum Leap.

It was widely accepted that the Bohr model contained substantial flaws. Not only did it throw up the quirky quantum leap, but it took no account of special relativity, it failed to explain why the electron orbit did not decay due to synchrotron radiation but most important of all it failed to explain the nature of the quantization of angular momentum*. The fact is that the assumption that angular momentum is quantized lacks any mathematical rigour, the assumption is arbitrary and expedient and fails to address the underlying question as to why and how such quantization occurs but merely asserts that if we make the assumption then the numbers seem to fit. Nevertheless, and despite this, the Bohr assumption has continued to be accepted and forms an integral part of every theory which has come along since.

In a paper published in 1905 Einstein had shown that light, which had hitherto been considered a wave, was in fact a particle [7]. In an effort to explain quantization the French mathematician Louis de Broglie turned this idea on its head and suggested that perhaps the electron was not a particle but should be considered as a wave instead. He calculated the wavelength of the electron, dividing Planck's constant by the electron's linear momentum and found that when he did so the orbital path of base energy state contained one wavelength; that of the second energy state contained two wavelengths and so on, in what appeared at first site to be a series of harmonics†.

On any other scale the wavelength of an object in orbit is associated with the orbital path length or circumference of the orbit and can be derived as a result of dividing the angular momentum of the orbiting object by its linear momentum. De Broglie instead chooses to associate the wavelength of the particle with the value of Planck's constant divided by the linear momentum, while at the same time assuming that the angular momentum of the particle was an integer multiple of Planck's constant. In choosing to substitute Planck's constant in this way instead of the angular momentum when calculating the wavelength, what de Broglie is doing is to coerce the wavelength of the electron to be an integer fraction of the orbital path length. Viewed in this light de Broglie's contribution can be seen as less of an insight and more of a contrivance.

If you were to observe an object in orbit, say a moon orbiting Jupiter or the proverbial conker‡ whirling on the end of a string, what you see is a sine wave. The orbiting object

*At first site it appears that the energy of the electron in the Bohr atom decreases with increasing energy level. However since the radius changes with energy level, the potential energy does also. When these two effects are combined, the energy levels increase with increasing energy level.

†In fact they are not harmonics of a single fundamental frequency, but instead each harmonic relates to a different base frequency and these two effects combine in such a way that they form a sub harmonic or inverse harmonic sequence

‡A conker is a horse chestnut on a string often used in a children's game

subtends a wave to an external observer of the form:

$$d = R\sin(\omega t) \tag{5}$$

or

$$d = R\sin(2\pi F t)\,. \tag{6}$$

where d is the displacement about some mean

For such a body we can easily calculate the orbital radius if we know the angular momentum and the linear momentum.

$$R = \frac{l}{p} = \frac{mvr}{mv}\,. \tag{7}$$

Furthermore we can identify the wavelength of such a wave with the orbital circumference which is simply.

$$\lambda = 2\pi R\,. \tag{8}$$

This is true for all orbiting objects no matter whether they are the size of a planet or the size of a conker.

By what rational then does de Broglie identify the wavelength of the orbiting electron, not with the angular momentum in this way, but with Planck's constant, which he believes, according to Bohr's assumption, to be an integer fraction of the angular momentum?

The alignment of wavelength with Planck's constant in this way cannot be justified either mathematically or mechanically. It is a contrivance which leads to the idea that there is some sort of wavelength which is an integer fraction of the orbital circumference. This is no miraculous discovery, not an insight into the workings of the atom, but an artificial device which reinforces and sustains the Bohr assumption without any basis in mechanics.

Other later models, such as that of Schrödinger, are based directly on the work of de Broglie and therefore inherently follow Bohr's assumption, up to and including the currently proposed Standard Model. Having been adopted by Bohr, later theorists simply continued with this working assumption and incorporated it into all subsequent models for the atom, without ever bothering to go back and justify it, until now it has become an item of received wisdom and an article of faith.

The trouble with all of these models is that the assumption proposed by Nicholson and adopted by Bohr is not based on finding any mechanism that leads to angular momentum being quantized in this way. The assumption was simply expedient — it just happens to give the values for the absorption and emission spectra of the hydrogen atom which match those of the Rydberg formula.

The year 1905 was an eventful one for Albert Einstein. In that year, he not only published his paper on the discrete nature of the photon but he also published two further seminal works as well as submitting his Ph.D. thesis. The most famous of his other papers concerned the dynamics of moving bodies [8]. This is the paper whose later editions contained the equation $e = mc^2$. The paper was based on a thought experiment and concerned the perception of time, distance and mass as experienced by two observers, one a stationary observer and one moving relative to the stationary observer at speeds approaching that of light.

What Einstein showed is that time elapses more slowly for a moving observer, that distances measured by a moving observer are foreshortened relative to those same distances measured by a stationary observer and that a stationary observer's perception of the mass of a moving object is that it has increased. All three effects occur to the same extent and are governed by a factor γ (Gamma). The time between two events observed by the stationary observer as time t is seen by the moving observer as time $T = t/\gamma$. Similarly the distance between two point measured by the stationary observer as distance d is seen by the moving observer as distance $D = d/\gamma$. As far as the stationary observer is concerned the mass of the moving object is seen to increase by this same factor γ.

Gamma is referred to as the Lorentz factor and is given by the formula:

$$\gamma = \frac{c}{\sqrt{c^2 - v^2}} = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}\,. \tag{9}$$

Both observers agree on their relative velocity but go about calculating it in different ways. For the stationary observer the velocity of the moving observer is the distance travelled divided by the time taken as measured in his stationary domain. For the stationary observer the velocity is:-

$$v = \frac{d}{t}\,. \tag{10}$$

For the moving observer the distance as measured in his own domain is foreshortened by the factor Gamma, but the time taken to cover that distance reduced by the same factor Gamma.

$$v = \frac{D}{T} = \frac{\frac{d}{\gamma}}{\frac{t}{\gamma}} = \frac{d}{t}\,. \tag{11}$$

There is a great deal of experimental evidence to support Einstein's Special Theory. One of the more convincing experiments was carried out at CERN in 1977 and involved measuring the lifetimes of particles called muons in an apparatus called the muon storage ring [9]. The muon is an atomic particle which carries an electric charge, much like an electron, only more massive. It has a short lifetime of around 2.2 microseconds before it decays into an electron and two neutrinos.

In the experiment muons are injected into a 14m diameter ring at a speed close to that of light, in fact at 99.94% of the speed of light where Gamma has a value of around 29.33. The muons, which should normally live for 2.2 microseconds, were seen to have an average lifetime of 64.5 microseconds; that is the lifetime of the muon was increased

by a factor Gamma. This comes about because the processes which take place inside the muon and which eventually lead to its decay are taking place in an environment which is moving relative to us at 99.94% of the speed of light and in which time, relative to us, is running 29.33 times slower. Hence the muon, in its own domain, still has a lifetime of 2.2 microseconds, it's just that to us, who are not moving, this appears as 64.5 microseconds.

Travelling at almost the speed of light a muon would normally be expected to cover a distance of 660 metres or roughly 7.5 times around the CERN ring during its 2.2 microsecond lifetime, but in fact the muons travelled almost 20,000 metres or 220 times around the ring. This is because distance in the domain of the muon is compressed so what we stationary observers see as being 20,000 metres the muon sees as being just 660 metres.

Both parties agree that during its lifetime the muon completes some 220 turns around the ring. We stationary observers see this as having taken place in some 64.5 microseconds, corresponding to a frequency of 3.4 MHz, while the muon sees these 220 turns as having been completed in just 2.2 microseconds, corresponding to a frequency of 100 Mhz. Hence for the muon and indeed all objects orbiting at close to light speed orbital frequency is multiplied by a factor Gamma relative to that of a stationary observer and it is this multiplication of orbital frequency which holds the key to the discrete energy levels of the atom.

As well as this effect on orbital frequency the muon ring experiment serves to show that considerations of special relativity can be applied to objects in orbit, this despite the fact that object in orbit are subject to a constant acceleration towards the orbital centre. However where the orbital velocity is constant, it is reasonable and correct to apply considerations of special relativity around the orbital path. In effect what we are doing is to resolve the orbital velocity into two components, one tangential component which has a constant velocity and one radial where there is a constant acceleration.

We have seen that speed is invariant with respect to relativity. Both the moving object and the stationary observer agree on their relative speed. This invariance of speed is central to the derivation of special relativity and so is deemed to be axiomatic. There is however one circumstance where it is reasonable to suggest that this need not be the case. For a stationary observer we normally require the use of two clocks in order to measure velocity; one at the point of departure and one at the point of arrival (at least conceptually). An object which is in orbit however returns once per cycle to its point of departure and so we can measure the orbital period of such an object with a single clock provided we do so over a complete orbit.

Thus for an object in orbit it is possible to define two velocity terms relating to the tangential or orbital velocity*. The first of these I have called the Actual Velocity and is simply the distance around the orbit divided by the orbital period as measured by the stationary observer. The second velocity term is the distance around the orbit as measured by the moving observer divided by the orbital period as measured by the stationary observer. Such a velocity term straddles or couples the two domains, that of the orbiting object and that of the stationary observer and so could sensibly be called the "Coupling Velocity" or possibly the "Relativistic Velocity". A simple calculation shows that the Relativistic Velocity is related to the Actual Velocity by the same factor Gamma an hence:

$$v_R = \frac{D}{t} = \frac{d}{t\gamma} = \frac{v}{\gamma} \,. \tag{12}$$

Thus far Relativistic Velocity is only a definition. However there is one set of circumstances where such a velocity term can indeed be justified and that is when dealing with the equations of motion relating to objects in orbit. It is considered here to be meaningful to use this Relativistic Velocity term when dealing with orbital velocities such as occur when calculating angular momentum, centripetal and centrifugal force and acceleration.

Nicholson had suggested that because Planck's constant has the units of angular momentum that it was somehow associated with the angular momentum of the orbiting electron. Here we take up that idea and suggest that the angular momentum of the orbiting electron is equal to Planck's constant, but reject his other idea that angular momentum is quantized. Instead we assume that orbital velocity is affected by relativity and use this to derive the equations of motion of the orbiting electron.

Planck's constant is then seen, not as a fundamental quantum of angular momentum but instead as providing a limiting value for angular momentum. The effect would not be significant at low velocities, but if the electron orbiting the hydrogen atom were to do so at close to light speed then:

$$l = \hbar = (m\gamma)\, r \left(\frac{c'}{\gamma} \right) . \tag{13}$$

where l is the angular momentum, $\hbar$ is Planck's constant, m is the mass of the electron, r is the orbital radius of the electron, c' is the orbital velocity of the electron and is very close to c, the speed of light.

Both the mass term and the velocity term are affected by relativity. The mass term because mass increases by factor Gamma as the object's velocity approaches the speed of light and in this case the velocity term is affected because we are dealing with an object in orbit and it is therefore appropriate

*In fact it is possible to define a further two velocity terms, the relativistic distance divided by the relativistic time and the actual distance divided by the relativistic time. The first of these is the invariant velocity discussed earlier. As a stationary observer we do not have any direct access to the moving clock and so these velocities can only be described mathematically and appear to have no physical significance.

to use Relativistic Velocity which is the Actual Velocity divided by Gamma. However since we are concerned here with an orbital velocity very close to the speed of light, to a first approximation we can substitute c for c' in Equation 13.

$$l = \hbar = (m\gamma)\, r \left(\frac{c}{\gamma}\right). \tag{14}$$

The two Gamma terms will cancel. The terms for rest mass, Planck's constant and the speed of light are all constants, which must therefore mean that the orbital radius is also a constant

$$R = \frac{\hbar}{mc}. \tag{15}$$

This not unfamiliar term is known as the Reduced Compton Wavelength although here it takes on a new and special significance as the characteristic radius at which an electron will orbit at or near light speed. This serves to explain why the orbiting electron does not emit synchrotron radiation. It does not do so because it is not driven to orbit the atomic nucleus by virtue of being accelerated by forces towards the orbital centre in the normal way, instead it is constrained to orbit at this radius by the limiting effect of Planck's constant. It is as if the electron is orbiting on a very hard surface from which it cannot depart and which it cannot penetrate. Equation 15 also means that there is no need to introduce the idea of a quantum leap or later equivalents. If the electron is constrained to always orbit at a fixed radius, then changes in energy level have to take place as a result of changes in orbital velocity, with no accompanying change of radius. Indeed this idea that the electron orbits at constant radius is a necessary condition for the electron to be considered objectively real.

Substituting Relativistic Velocity into the force balance equation that Bohr himself used, but at an orbital velocity very close to that of light yields another interesting result*

$$\frac{Kq^2}{\hbar c} = \frac{(m\gamma)}{r}\left(\frac{c}{\gamma}\right)^2. \tag{16}$$

Which combines with Equation 15 and simplifies to give:

$$\frac{Kq^2}{\hbar c} = \frac{1}{\gamma}. \tag{17}$$

Readers may be familiar with the term on the left of this equation which is known as the Fine Structure Constant often written as α (Alpha). So for the base energy state of the atom

$$\gamma = \frac{1}{\alpha}. \tag{18}$$

α has a value of $7.2973525698 \times 10^{-3}$

From this and Equation 9 we can easily calculate the corresponding orbital velocity and frequency as measured by the stationary observer.

$$\frac{v}{c} = \sqrt{1-\alpha^2} = 0.999973371. \tag{19}$$

The orbital velocity turns out to be 99.9973% of the speed of light c, thus vindicating the first approximation made in Equation 14 and the frequency (in the domain of the stationary observer)

$$\omega_1 = \frac{v}{R} = 7.76324511 \times 10^{20}. \tag{20}$$

The physicist Richard Feynman [10] once said of Alpha that:

"It has been a mystery ever since it was discovered more than fifty years ago, and all good theoretical physicists put this number up on their wall and worry about it. Immediately you would like to know where this number for a ***coupling***† *comes from: is it related to pi or perhaps to the base of natural logarithms? Nobody knows. It's one of the greatest damn mysteries of physics: a magic number that comes to us with no understanding by man. You might say the "hand of God" wrote that number, and "we don't know how He pushed his pencil." We know what kind of a dance to do experimentally to measure this number very accurately, but we don't know what kind of dance to do on the computer to make this number come out, without putting it in secretly!"*

Equation 18 effectively solves the mystery, providing an explanation for the physical significance of the Fine Structure Constant. It is seen simply as the ratio of two velocities, the Relativistic Velocity and the Actual Velocity of the orbiting electron. Since these two velocities share the same orbital period, it can also be seen as the ratio of two orbital path lengths, the one traversed at non-relativistic speeds to that traversed by the orbiting electron at near light speed. The Fine Structure Constant is seen to be dynamic in nature. Its value relies on the fact that the electron is in motion, orbiting at near light speed; it does so at a speed that is necessary to maintain structural equilibrium within the hydrogen atom, since it is only by travelling at this speed that the structural integrity of the atom can be maintained. In the world of the atom, where there is no friction and in the absence of any sort of external input, the atom remains stable and, unless disturbed in some way, the electron will continue in this state indefinitely. In this sense it defines the speed at which the electron has to travel in order to achieve a stable orbit.

So far we have only considered the lowest or base energy state of the atom. We have seen that one of the effects of relativity is to multiply frequency in the domain of a moving object by Gamma. The frequency in the domain of the

*Once again since the orbital velocity is very close to the speed of light we can, to a first approximation, substitute c as the Actual Velocity

†My emphasis — the term Coupling Velocity resonates with the idea of Alpha as a coupling constant.

electron which corresponds to this stable state is simply calculated by multiplying by Gamma — equivalent to dividing by Alpha – to give.

$$\Omega = \frac{\omega_1}{\gamma} = 1.06378925 \times 10^{23} . \quad (21)$$

But just as was the case with the observer and the cannon if there is a frequency Ω at which the atom is stable then frequencies of $n\Omega$ must also be stable for all n = integer which in turn means that there are stable states for all

$$\gamma_n = \frac{n}{\alpha} \quad (22)$$

and so

$$r_n = R = \frac{\hbar}{mc} \quad (23)$$

and

$$\frac{v_n}{c} = \sqrt{\frac{n^2 - \alpha^2}{n^2}} . \quad (24)$$

Equation 23 shows that the orbital radius remains the same for all energy levels, while Equation 24 describes the orbital velocity for the nth energy state*. Table 1 shows the resulting orbital velocities for the first 13 energy states and the theoretically infinite state of the hydrogen atom and as you might expect they match the absorption and emission spectra of the hydrogen atom perfectly.

During the 1930's and 40's Einstein and Bohr disagreed over the nature of reality, with Bohr arguing that the laws of physics were different on the scale of the atom and that as a consequence reality becomes subjective in nature. Particles are not considered to discrete point particles in the classical sense, but instead are considered to be nebulous wave-particles which manifest themselves as either particles or as waves when subjected to some sort of observing process. Einstein on the other hand took the view that reality had to be objective and that particles must therefore be discrete point particles having deterministic position and velocity.

In the end the debate was largely resolved by default. Bohr simply outlived Einstein and so his ideas prevailed and form the basis of today's Standard Model. Einstein is nowadays often described as being an old man, set in his ways and unable to accept the new ways of thinking. But this is to misconstrue Einstein's position, which was one of principle.

Einstein had argued that the laws of physics are the same for all reference frames, while Bohr reasoned that the laws of physics are different on the scale of the atom. Einstein was concerned with reference frames of comparable scale that were in motion with respect to one another but it is logical to extend his idea to reference frames of differing scales. If we start from this position and pursue the idea that particles are objectively real and that the laws of physics are the same independent of scale then it is necessary to question our current understanding of the laws of physics. They must be deficient in some way and it is necessary to find a way in which the laws must be modified to describe the atom but which does not affect our understanding on all other scales.

The idea of relativistic velocity postulated here does just that. It provides a model for the structure and dynamics of the hydrogen atom which is consistent with particles which are objectively real. At the same time it does what all previous models have failed to do and provides a mechanism to explain exactly why the energy levels of the atom are quantized without the need of resorting to arbitrary assumptions. The idea of a Relativistic Velocity or Coupling Velocity, a velocity term which is affected by relativity, solves all of the problems that faced Niels Bohr with his model and produces a model for the hydrogen atom which matches the emission and absorption spectra of the atom.

Here quantization takes place with respect to the variable Gamma as the orbital velocity of the electron gets ever closer to the speed of light with increasing energy level, and not with respect to angular momentum as postulated by Bohr. Angular momentum for the orbiting electron remains substantially constant and equal to Planck's constant over all of its energy levels as the orbital velocity varies from 99.99733% of c for the base energy state upwards as energy levels increase, although never quite achieving the theoretical limit of 100%, while Gamma is constrained to take on values which are integer multiples of a base value, that value being the reciprocal of the Fine Structure Constant. Planck's constant takes on a new and special significance, not as the quantum of angular momentum of the existing models, but as a lower limit for angular momentum below which it cannot exist.

The orbital radius of the electron remains substantially constant irrespective of the energy level of the atom, a necessary condition for an objectively real electron, and so transitions from one energy state to another take place without the need to introduce the idea of discontinuity of position, inherent in the Bohr model, or its equivalent probability density functions and wave particle duality found in other more recent models. Such transitions are easily explained as simple changes in the orbital velocity of the electron over a dynamic range which lies very close to the speed of light. With no changes in orbital radius, changes in energy level involve no change in potential energy, only the kinetic energy of the orbiting electron changes between energy states.

Thus the morphology of the atom remains substantially unaltered for all energy levels. This is consistent with the atom having the same physical and chemical properties irrespective of energy level. The Bohr model, and indeed the standard model, would have us believe that the morphology of the atom changes substantially with energy level, with the orbital radius increasing as the square of the energy level with no theoretical upper limit. Such changes are difficult to rec-

*Notice that since the orbital radius remains substantially the same for all energy levels, there is no change in potential energy between the various different energy levels, only a change in kinetic energy.

n	v_n/c	$1/\gamma_n$	Energy eV	ΔEnergy	eV
1	0.999973371	0.007297559	7.76324511E+20	255485.925	13.607
2	0.999993343	0.003648853	7.76340016E+20	255496.130	3.402
3	0.999997041	0.002432577	7.76342887E+20	255498.020	1.512
4	0.999998336	0.001824435	7.76343892E+20	255498.682	0.850
5	0.999998935	0.001459549	7.76344357E+20	255498.988	0.544
6	0.999999260	0.001216291	7.76344610E+20	255499.154	0.378
7	0.999999457	0.001042536	7.76344762E+20	255499.255	0.278
8	0.999999584	0.000912219	7.76344861E+20	255499.320	0.213
9	0.999999671	0.000810861	7.76344929E+20	255499.364	0.168
10	0.999999734	0.000729775	7.76344977E+20	255499.396	0.136
11	0.999999780	0.000663432	7.76345013E+20	255499.420	0.112
12	0.999999815	0.000608146	7.76345040E+20	255499.438	0.094
13	0.999999842	0.000561366	7.76345061E+20	255499.452	0.081
∞	1.000000000	0.000000000	7.76345184E+20	255499.532	0.000

Table 1:

oncile with an atom who's physical and chemical properties remain the same for all energy levels.

The model explains all of the shortcomings found in the Bohr model, the absence of orbital decay due to synchrotron radiation and the need for a quantum leap. Bohr had ignored the effects of special relativity on the energy levels of the atom, even though they should have been small but significant at the velocities predicted by his model. Here they are fully integrated into the model.

The model sheds a new light on the nature of the wave particle duality. The electron is seen as a point particle in the classical sense, having deterministic position and velocity*. Electrons are thus objectively real. The electron has wavelike properties, but these derive from the orbital motion of an objectively real particle. The waves are seen as the projection of the circular orbit of the objectively real electron onto an external observer, in much the same way that we can describe the orbit of the moons of distant planets as having a wavelike nature. There is no need to invent the ether or what has more recently passed for the ether, the so called fabric of space time, as a medium in which these waves exist. In the final analysis where vacuum contains absolutely nothing, there is nothing to wave except the particle and that is precisely what the model provides.

The introduction of Relativistic Velocity has another major implication. It extends the laws of physics down to the scale of the atom and possibly beyond. With its introduction the same set of physical laws extends from a scale of approximately 10^{-20} m to 10^{20} m thus doing away with the notion that a different set of physical law applies on the scale of the atom. It is quite likely therefore that a single set of physical laws exists for all scales and throughout the universe.

Finally it provides a simple mechanical explanation for the existence and the value of the hitherto mysterious Fine Structure Constant.

Appendix 1 Derivation of Centripetal Acceleration under relativistic conditions

The idea that orbital velocity is affected by relativity is central to the theory presented here, so it is perhaps worthwhile examining this idea in a little more detail. Before doing so however it is necessary to restate that the use of Special Relativity in dealing with objects which have constant orbital velocity is entirely appropriate, this despite the fact that such objects are subject to acceleration. The velocity of an object which is in orbit can be considered as having two components, a tangential component and a radial component. For constant orbital velocity, the tangential component is itself constant and therefore can be dealt with using Special Relativity which affects the time and distance measured along the orbital path. Direct evidence to support this comes in the form of the Muon ring experiment described earlier.

Such an orbiting object is subject to constant acceleration towards the orbital centre and it is this acceleration which in effect maintains the circular path. Conventional wisdom has it that this centripetal acceleration is not affected by relativity, since it acts in a direction which is normal to the velocity of the object. Here it is argued that this cannot be the case since the distances involved in calculating centripetal acceleration derive directly from the distances travelled around the orbital path and that these distances are themselves affected by relativity. It can then be shown that this is equivalent to substituting Relativistic Velocity in place of Actual Velocity in the standard formula for calculating centripetal acceleration.

Einstein showed that objects which are travelling at close

*This is not to say that uncertainty does not exist, it does, but it is seen as a practical issue of measurement when the scale of the measurement tools is similar to that of the measured object and not as being an intrinsic property of the particle.

to light speed are affected in three ways, time in the domain of the moving observer advances at a slower rate than it does for a stationary observer, distance for the moving object is foreshortened in the direction of travel relative to that same distance as measured by the stationary observer. The mass of a moving object appears increased as far as the stationary observer is concerned. All three effects occur to the same extent by the factor Gamma (γ). Gamma is named after the Dutch physicist Hendrik Antoon Lorentz (1853 — 1928). Gamma is given by the formula

$$\gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} . \tag{25}$$

Examination of the effect of relativity on an object moving at close to the speed of light however reveals that both time and distance are scaled by a factor $1/\gamma$ and so from Equation 25

$$\frac{1}{\gamma} = \sqrt{1 - \frac{v^2}{c^2}} . \tag{26}$$

It can be seen that this is the equation of a circle, more specifically a quadrant of a unit circle, since v is constrained to lie between 0 and c as shown in Figure 1.

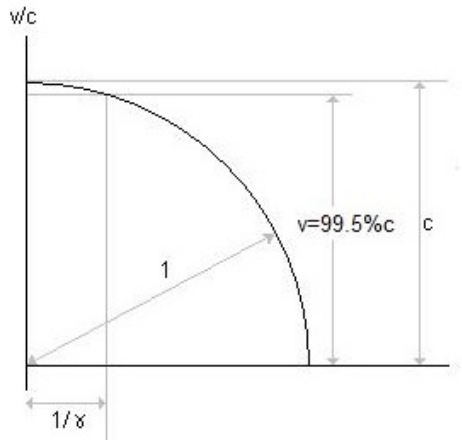

Fig. 1:

If the object under consideration is in circular orbit, then this quadrant can be superimposed on the orbital path to form a hemisphere. Objects orbiting at non-relativistic speeds see the path length around the orbit as being equal in length to the equator, while objects orbiting at higher speeds follow a path length described by a line of latitude on the hemisphere. An object orbiting at the theoretical maximum speed of light would then be pirouetting at the pole. We can consider the length of the orbital path as being represented by the line of latitude formed by a slicing plane which cuts through the hemisphere parallel to the equatorial plane. In Figure this is at approximately 15% of the speed of light c and so the orbital path length is just a little less than the equatorial path length, around 99%.

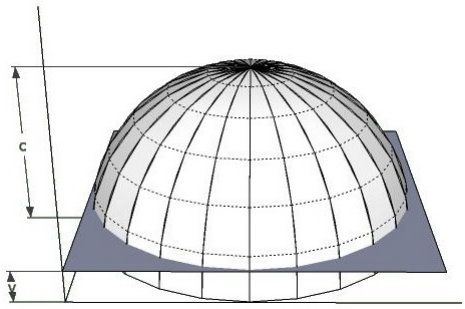

Fig. 2:

In Figure 3 the orbital velocity is approximately 80% of the speed of light and so the orbital path length as seen by the moving object is approximately 60% that for an object moving at non-relativistic speed

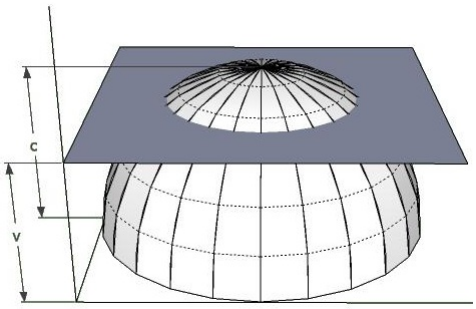

Fig. 3:

In Figure 4 the orbital velocity is around 98% of the speed of light and the corresponding orbital path length is approximately 20% of that for non-relativistic motion.

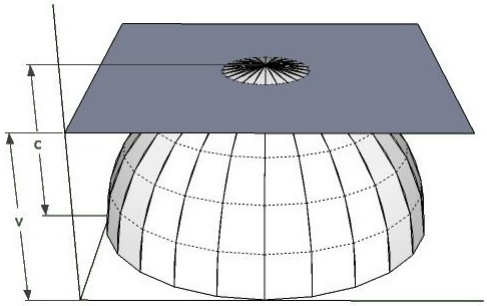

Fig. 4:

This hemispheric model of the motion of an orbiting object is useful because it allows us to visualise the orbital path

length as being foreshortened by relativity while at the same time the radius of the orbit is unaffected by relativity. The orbital geometry is non-Euclidean and in reality all takes place in just one plane. The introduction of this third dimension is just a device to allow us to visualise what is going on. The orbiting object sees the distance it travels around one orbit as being reduced by a factor Gamma, but nevertheless sees the orbital radius as being unaffected by relativity since this is at right angles to the direction of travel. Thus we can represent the radius of the orbit as being the distance from a point on the relativistic orbit to the centre of the hemisphere.

The term Actual Velocity has been adopted to describe the velocity of the orbiting object as seen by a stationary observer. This is easily calculated as the circumference of the orbital path, the equator of the hemisphere (d), divided by the orbital period (t), both measured by the stationary observer.

The theory postulates that there is a velocity term which is affected by Gamma. This is termed the Relativistic Velocity, but only becomes significant when the Actual Velocity is close to the speed of light. This velocity term can be calculated by taking the foreshortened distance around the line of latitude, which represents the orbital path as seen by the moving observer, divided by the orbital period as measured by a stationary observer. The foreshortened distance around the orbit is calculated as d/γ and the orbital period remains the same as for Actual Velocity (t) and hence this Relativistic Velocity is then easily calculated as $v_R = d/t\gamma$.

We can use this term directly in calculating the angular momentum of the orbiting object. This is simply a restatement of the argument used earlier. Angular momentum is the product of the mass, the velocity and the radius of an orbiting point object. However the mass of the object is affected by relativity, appearing to increase the mass by a factor Gamma (γ) and so:

$$l = (m\gamma)\, r \left(\frac{v_R}{\gamma} \right) . \tag{27}$$

However since for Gamma to take on a significant value v_R must be very close to c, the speed of light and so we can substitute c for v_R. Also since the angular momentum of an electron in orbit around an atomic nucleus is given by Planck's constant we can substitute this for l in Equation 27 to give:

$$l = \hbar = mcr . \tag{28}$$

In effect we are simply substituting Relativistic Velocity for Actual Velocity in the standard textbook formula for calculating angular momentum. This is recognising that the orbital velocity is the distance around the orbit as measured by the moving object divided by the orbital period as measured by a stationary observer.

We can of course use this same argument to substitute Relativistic Velocity for Actual Velocity in the formula for centripetal acceleration and hence derive expressions for centripetal and centrifugal forces. However in the case of centripetal acceleration it is also useful to derive an expression for the relativistic case from first principles.

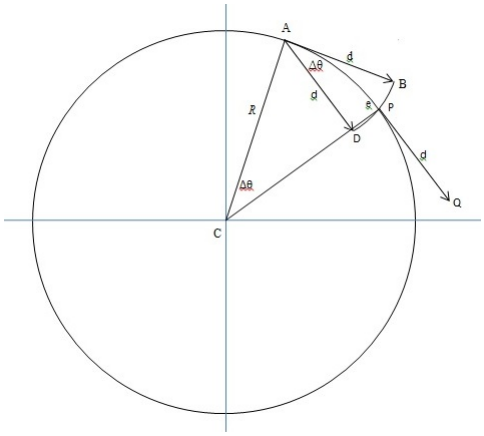

Fig. 5:

The formula for centripetal force was first derived by Christian Huygens in 1659 and describes a constant force acting on a body in circular motion towards the centre of the circle. When combined with Newton's second law this leads to the idea that a body in circular motion is subject to a constant acceleration towards the centre called centripetal acceleration.

It is customary when deriving the formula for centripetal acceleration to use velocity vectors directly. Here we take a slightly different approach and use the distance vectors instead. This is because in the proposed theory only the distance component of velocity is affected by relativity and not the time component. In other respects the derivation is the same as that found in many standard texts.

Consider an object in orbit around a point C at radius R. At a particular instant t the object is at point A and some short interval of time later Δt it is at point P, having moved through an angle subtended at the centre of the circle of $\Delta\theta$.

The vector representing the distance moved in time Δt is AB and has length d and is tangential to the circle, hence CAB is a right angle. At $t + \Delta t$ the object is at P and has a distance vector PQ, also of length d. We can translate the vector PQ to A forming AD. The vector BD then represents the distance moved towards the centre of the circle in time Δt. Note that for as $\Delta\theta$ tends to 0 the line BD tends to a straight line.

Then

$$d = R\Delta\theta . \tag{29}$$

Since APC and ABD are similar triangles (for small $\Delta\theta$)

$$e = d\Delta\theta \tag{30}$$

and the acceleration towards the centre of the circle is

$$a = \frac{e}{\Delta t^2} . \tag{31}$$

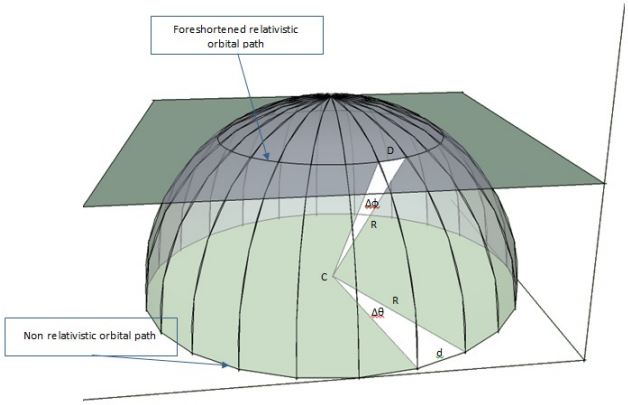

Fig. 6:

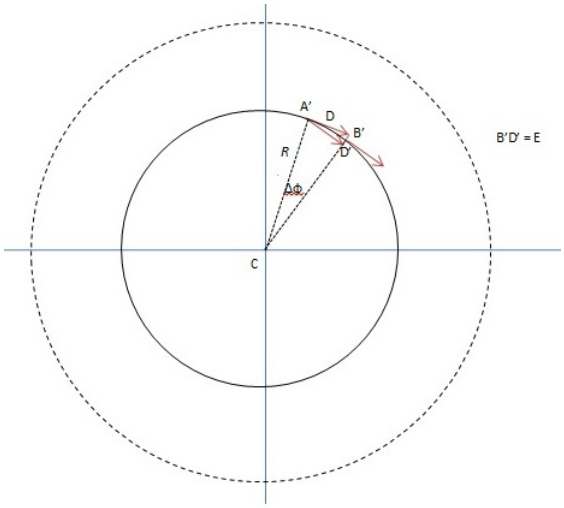

Fig. 7:

Therefore

$$a = \frac{R\Delta\theta^2}{\Delta t^2} \,. \tag{32}$$

Multiplying both top and bottom by R gives

$$a = \frac{R^2\Delta\theta^2}{R\Delta t^2} \,. \tag{33}$$

But since

$$v = \frac{d}{\Delta t} = \frac{R\Delta\theta}{\Delta t} \,. \tag{34}$$

Then

$$a = \frac{v^2}{R} \,. \tag{35}$$

When we take into consideration the effects of special relativity, the situation becomes a little more complicated. Although the orbital path is foreshortened, as represented by the line of latitude in Figure 6, and hence the circumference of this circle is reduced by a factor Gamma, the radius of the circle is not affected and remains the same as that for the equatorial orbital path.

Figure 6 attempts to show this by introducing a third dimension and using the hemispherical representation developed above. In reality however the radius and the orbital path are co-planar. It can be seen from Figure 6 that the angle subtended by a short segment of the circumference is less for the relativistic path than for the non-relativistic path. From Figure 6 it is evident that

$$\Delta\phi = \frac{\Delta\theta}{\gamma} \tag{36}$$

and

$$R\Delta\phi = \frac{R\Delta\theta}{\gamma} \,. \tag{37}$$

Figure 7 shows the foreshortened orbital path in plan view. The dashed circle represents the non-relativistic orbital path while the radii are shown dotted to indicate that they are not to scale in this representation.

The distance travelled during time Δt is foreshortened by relativity, instead of travelling a distance AB the object only travels a distance A'B'=D in Figure 7.

$$D = R\Delta\phi \,. \tag{38}$$

Once again the triangles CA'B' and A'B'D' are similar and so the distance travelled towards the centre of the orbit E is

$$E = D\Delta\phi \,. \tag{39}$$

Once again the triangles CA'B' and A'B'D' are similar and so the distance travelled towards the centre of the orbit E is

$$A = \frac{E}{\Delta t^2} \,. \tag{40}$$

Which is also

$$A = \frac{R\Delta\phi^2}{\Delta t^2} \tag{41}$$

Again we can multiply both denominator and numerator by R to give

$$A = \frac{R^2\Delta\phi^2}{R\Delta t^2} \,. \tag{42}$$

Which gives

$$A = \frac{R^2\Delta\theta^2}{R\Delta t^2\gamma^2} \tag{43}$$

and so

$$A = \frac{v^2}{R\gamma^2} \,. \tag{44}$$

Equation 44 represents a more general case for calculating centripetal acceleration. When the orbital velocity is low, under non-relativistic conditions, the value of Gamma is unity

and the formula can be simplified to the more familiar one shown in Equation 35. Effectively therefore the formula for centripetal acceleration under relativity substitutes Relativistic Velocity for Actual Velocity in the standard textbook formula.

It is the geometry of the triangle AB'D' which lies at the heart of the argument. Here it is argued that the length B'D' is affected by relativity even though it is measured in a direction at right angles to the direction of travel. This comes about because the lengths of the two sides AB' and AD' are both themselves affected by relativity and the triangle must have geometric integrity and so B'D' must also be scaled by relativity. If it was not then the triangle AB'D' would be a very strange triangle indeed. It would have to be an isosceles triangle in which the third side could be longer than the sum of the two other sides. The direction of the vectors AB' and AD' could not be preserved. Even in non-Euclidian geometry such a triangle would not be possible and so B'D' must be scaled by Gamma.

The measurement of time on the other hand can only take place in the domain of the observer, so the moving observer sees his time in his own domain and the stationary observer sees time in his domain. The two domains are related by a factor Gamma, but from the point of view of direct measurement this is a theoretical connection. In other words the stationary observer has no direct access to the moving clock and, vice versa, the moving observer has no direct access to the stationary clock.

Appendix 2 An Analytical Method for calculating Actual Velocity

A more analytical approach for calculating the value for c' can be found without the first approximation used above:

The equation for the value of gamma

$$\gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} . \tag{45}$$

From which

$$v = c\sqrt{\frac{\gamma^2 - 1}{\gamma^2}} . \tag{46}$$

Substituting this into the force balance equation gives

$$\frac{m_0 c^2(\gamma^2 - 1)}{R\gamma^3} = \frac{Kq^2}{R^2} . \tag{47}$$

Recognising that $\hbar = m_0 Rc$ and simplifying gives

$$\frac{\gamma^2 - 1}{\gamma^3} = \frac{Kq^2}{\hbar c} . \tag{48}$$

The term on the right hand side is the Fine Structure Constant which is denoted by α. Substituting and rearranging gives the following equation for γ.

$$\alpha\gamma^3 - \gamma^2 + 1 = 0 . \tag{49}$$

The numerical value for α^* is $7.2973525698 \times 10^{-3}$. Substituting this and calculating the three roots gives:

$\gamma = 137.028700944403$

$\gamma = -0.996384222264$

$\gamma = 1.0036823521665$

Only the first of these three values is significant. This cubic equation gives a more precise value for Gamma. By recognizing that v is very close to c in the force balance equation the value of Gamma can be calculated as:

Substituting in the equation for γ gives a value for v:

$$v = c\sqrt{\frac{\gamma^2 - 1}{\gamma^2}} = 0.999973371c . \tag{50}$$

v is the Actual Velocity of the electron around its orbit and as can be seen it is very close to c, the velocity of light, being some 99.9973371% of c, which is in agreement with the method of first approximation to the first 8 significant figures.

Appendix 3 The Rydberg Formula

Joseph Jakob Balmer (1825–1898) was a Swiss mathematician and numerologist who, after his studies in Germany, took up a post teaching mathematics at a girls' school in Basel. A colleague in Basel suggested that he take a look at the spectral lines of hydrogen to see if he could find a mathematical relationship between them. Eventually Balmer did find a common factor† $h = 3.6456 \times 10^{-7}$ which led him to a formula for the wavelength of the various spectral lines.

$$\lambda = \frac{hm^2}{m^2 - 4} , \tag{51}$$

where m is an integer with value 3 or higher.

Balmer originally matched his formula for $m = 3, 4, 5, 6$ and based on this he predicted an absorption line for $m = 7$. Balmer's seventh line was subsequently found to match a new line in the hydrogen spectrum that had been discovered by Ångström.

Balmer's formula dealt with a particular set of spectral lines in the hydrogen atom and was later found to be a special case of a more general result which was formulated by the Swedish physicist Johannes Rydberg.

$$\frac{1}{\lambda} = R_H\left(\frac{1}{n_1^2} - \frac{1}{n_2^2}\right) , \tag{52}$$

where λ is the wavelength of the spectral line, R_H is the Rydberg constant for hydrogen, n_1 and n_2 are integers and $n_1 < n_2$.

By setting n_1 to 1 and allowing n_2 to take on values of $2, 3, 4 \ldots \infty$ the lines take in a series of values known as the Lyman series. Balmer's series is obtained by setting $n_1 = 2$ and allowing n_2 to take on values of $3, 4, 5 \ldots \infty$. Similarly for other values of n_1 series of spectral lines have been named according to the person who first discovered them and so:

*CODATA - http://physics.nist.gov/cgi-bin/cuu/Value?alph

†h here is not to be confused with Planck's constant.

n_1	n_2	Series
1	$2 \ldots \infty$	Lyman series
2	$3 \ldots \infty$	Balmer series
3	$4 \ldots \infty$	Paschen series
4	$5 \ldots \infty$	Brackett series
5	$6 \ldots \infty$	Pfund series
6	$7 \ldots \infty$	Humfreys series

Other series beyond these do exist, but they are not named.

By substituting different values for R, it was found that Rydberg's formula worked for all so called *hydrogenic*[*] atoms.

The value of R_H can be found by considering the case where $n_1 = 1$ and $n_2 = \infty$, a condition which represents the maximum possible change in energy level within the hydrogen atom. R_H is then the wavelength of the absorption line associated with such an energy change and was calculated to have a value of 1.097×10^7

This was subsequently found to be given by the formula:

$$R_H = \frac{1}{4\pi} \frac{m_0 c \alpha^2}{\hbar} . \tag{53}$$

The highest possible energy level for the atom occurs when n, the energy level, equals the theoretical value of infinity. The corresponding value for the Actual Velocity would then be c, the speed of light.

The equation for the energy of an orbiting body of mass m with velocity v is easily obtained in any standard text and is given by:

$$e = \frac{1}{2} m v^2 . \tag{54}$$

If we assume that the electron is orbiting at near light speed then the maximum possible energy[†] of an electron orbiting the hydrogen nucleus where the orbital velocity has a theoretical value of c, the speed of light and the mass of the electron is m_0 is

$$e = \frac{1}{2} m_0 c^2 . \tag{55}$$

The energy potential for a hydrogen atom in any arbitrary energy state n is the difference between this maximum energy value and the energy of the nth state

$$e_n = \frac{1}{2} m_0 c^2 - \frac{1}{2} m_0 v_n^2 = \frac{1}{2} m_0 (c^2 - v_n^2) . \tag{56}$$

[*]A hydrogenic atom is one which is ionized such that it has only one orbiting electron. In theory, at least, any atom can be ionized so as to become hydrogenic.

[†]Note that the electron is orbiting at the same radius for all energy levels, the potential energy of the electron therefore remains the same and all changes in energy level which are then associated with changes in kinetic energy and hence with the velocity of the electron.

We saw earlier that gamma could be expressed in terms of c, the velocity of light and v, the Actual Velocity using Einstein's equation for special relativity and that $\gamma_n = n\gamma_0$

$$\gamma_n = \frac{c}{\sqrt{c^2 - v_n^2}} . \tag{57}$$

This is easily rearranged to give an expression for $c^2 - v^2$

$$c^2 - v_n^2 = \frac{c^2}{\gamma_n^2} \tag{58}$$

In the base energy state $n = 0$ and $\gamma_0 = 1/\alpha$

$$c^2 - v_0^2 = c^2 \alpha^2 \tag{59}$$

Hence the maximum energy potential for the atom is

$$e_p = \frac{1}{2} m_0 c^2 \alpha^2 . \tag{60}$$

Substituting numerical values for m_0, c and α gives the maximum energy potential of the atom as
$e_p = 2.18009839 \times 10^1 8$ Joules
or
$e_p = 13.6071$ eV.

The energy potential for any arbitrary energy level n is given by

$$e_{pn} = \frac{1}{2} \frac{m_0 c^2 \alpha^2}{n^2} . \tag{61}$$

Hence the difference between any two energy levels n and m is

$$e_{n,m} = \frac{1}{2} m_0 c^2 \alpha^2 \left(\frac{1}{n^2} - \frac{1}{m^2} \right) . \tag{62}$$

and the difference in orbital frequency is

$$\omega_{n,m} = \frac{1}{2} \frac{m_0 c \alpha^2}{\hbar} \left(\frac{1}{n^2} - \frac{1}{m^2} \right) \tag{63}$$

This can be expressed in terms of wavelength, similar to the Rydberg formula, by dividing both sides by 2π to give

$$\frac{1}{\lambda_{n,m}} = \frac{1}{4\pi} \frac{m_0 c \alpha^2}{\hbar} \left(\frac{1}{n^2} - \frac{1}{m^2} \right) \tag{64}$$

and

$$R_H = \frac{1}{4\pi} \frac{m_0 c \alpha^2}{\hbar} . \tag{65}$$

Submitted on: October 8, 2012 / Accepted on: October 11, 2012

References

1. Heilbron J.L. Historical Studies in the Theory of Atomic Structure. *Arno Press*, 1981.
2. Nyquist H. Certain Topics in Telegraph Transmission Theory. *Trans. AIEE*, Apr 1928, v.47, 617–644, .
3. Shannon C.E. Communication in the Presence of Noise. *Proc. Institute of Radio Engineers*, Jan 1949, v.37, no.1, 10–21.

4. Bracewell R.N. The Fourier Transform and Its Applications (revised ed.). *McGraw-Hill*, 1st ed., 1965, 2nd ed. 1978.
5. Bohr N. On the Constitution of Atoms and Molecules. Part I. *Philosophical Magazine*, 1913, v.26, 1–24.
6. Nicholson J.W. A structural Theory of the Chemical Elements. *Philosophical Magazine*, 1911, v.22, 864–889.
7. Einstein A. Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt (On a Heuristic Viewpoint Concerning the Production and Transformation of Light). *Annalen der Physik*, 1905, Bd.17(6), 132–134.
8. Einstein A. Zur Elektrodynamik bewegter Körper (On the Electrodynamics of Moving Bodies). *Annalen der Physik*, 1905, Bd.17(10), 891–921.
9. Bailey H., Borer K., Combley F., Drumm H., Krienen F., Lange F., Picasso E., von Ruden W., Farley F.J.M., Field J.H., Flegel W. & Hattersley P.M. Measurements of relativistic time dilatation for positive and negative muons in a circular orbit. *Nature*, 1977, v.268(5618), 301–305.
10. Feynman R.P. QED: The Strange Theory of Light and Matter. Princeton University Press, 1985, p.129.

The Elastodynamics of the Spacetime Continuum as a Framework for Strained Spacetime

Pierre A. Millette

University of Ottawa (alumnus), K4A 2C3 747, Ottawa, Canada. E-mail: PierreAMillette@alumni.uottawa.ca

We derive the elastodynamics of the spacetime continuum by applying continuum mechanical results to strained spacetime. Based on this model, a stress-strain relation is derived for the spacetime continuum. From the kinematic relations and the equilibrium dynamic equation of the spacetime continuum, we derive a series of wave equations: the displacement, dilatational, rotational and strain wave equations. Hence energy propagates in the spacetime continuum as wave-like deformations which can be decomposed into dilatations and distortions. Dilatations involve an invariant change in volume of the spacetime continuum which is the source of the associated rest-mass energy density of the deformation, while distortions correspond to a change of shape of the spacetime continuum without a change in volume and are thus massless. The deformations propagate in the continuum by longitudinal and transverse wave displacements. This is somewhat reminiscent of wave-particle duality, with the transverse mode corresponding to the wave aspects and the longitudinal mode corresponding to the particle aspects. A continuity equation for deformations of the spacetime continuum is derived, where the gradient of the massive volume dilatation acts as a source term. The nature of the spacetime continuum volume force and the inhomogeneous wave equations need further investigation.

1 Introduction

Strained spacetime has been explored recently by Millette [1] from a continuum mechanical and general relativistic perspective, and by Tartaglia *et al* in the cosmological context, as an extension of the spacetime Lagrangian, to obtain a generalized Einstein equation [2, 3].

As shown in [1], the applied stresses from the energy-momentum stress tensor result in strains in the spacetime continuum. The presence of strains as a result of applied stresses is an expected continuum mechanical result. The strains result in a deformation of the continuum which can be modeled as a change in the underlying geometry of the continuum. The geometry of the spacetime continuum of General Relativity resulting from the energy-momentum stress tensor can thus be seen as a representation of the deformation of the spacetime continuum resulting from the strains generated by the energy-momentum stress tensor.

In this paper, we examine in greater details the elastodynamics of the spacetime continuum as a framework for describing strained spacetime.

2 Elastodynamics of the Spacetime Continuum

2.1 Model of the Elastodynamics of the Spacetime Continuum

The spacetime continuum (*STC*) is modelled as a four-dimensional differentiable manifold endowed with a metric $g_{\mu\nu}$. It is a continuum that can undergo deformations and support the propagation of such deformations. A continuum that is deformed is strained.

An infinitesimal element of the unstrained continuum is characterized by a four-vector x^μ, where $\mu = 0, 1, 2, 3$. The time coordinate is $x^0 \equiv ct$.

A *deformation* of the spacetime continuum corresponds to a state of the *STC* in which its infinitesimal elements are displaced from their unstrained position. Under deformation, the infinitesimal element x^μ is displaced to a new position $x^\mu + u^\mu$, where u^μ is the displacement of the infinitesimal element from its unstrained position x^μ.

The spacetime continuum is approximated by a deformable linear elastic medium that obeys Hooke's law. For a general anisotropic continuum in four dimensions [4, see pp. 50–53],

$$E^{\mu\nu\alpha\beta}\varepsilon_{\alpha\beta} = T^{\mu\nu} \tag{1}$$

where $\varepsilon_{\alpha\beta}$ is the strain tensor, $T^{\mu\nu}$ is the energy-momentum stress tensor, and $E^{\mu\nu\alpha\beta}$ is the elastic moduli tensor.

The spacetime continuum is further assumed to be isotropic and homogeneous. This assumption is in agreement with the conservation laws of energy-momentum and angular momentum as expressed by Noether's theorem [5, see pp. 23–30]. For an isotropic medium, the elastic moduli tensor simplifies to [4]:

$$E^{\mu\nu\alpha\beta} = \lambda_0(g^{\mu\nu}g^{\alpha\beta}) + \mu_0(g^{\mu\alpha}g^{\nu\beta} + g^{\mu\beta}g^{\nu\alpha}) \tag{2}$$

where λ_0 and μ_0 are the Lamé elastic constants of the spacetime continuum. μ_0 is the shear modulus (the resistance of the continuum to *distortions*) and λ_0 is expressed in terms of κ_0, the bulk modulus (the resistance of the continuum to *dilatations*) according to

$$\lambda_0 = \kappa_0 - \mu_0/2 \tag{3}$$

in a four-dimensional continuum. A *dilatation* corresponds to a change of volume of the spacetime continuum without a change of shape while a *distortion* corresponds to a change of shape of the spacetime continuum without a change in volume.

2.2 Stress-Strain Relation of the Spacetime Continuum

Substituting Eq.(2) into Eq.(1), we obtain the stress-strain relation for an isotropic and homogeneous spacetime continuum

$$2\mu_0\varepsilon^{\mu\nu} + \lambda_0 g^{\mu\nu}\varepsilon = T^{\mu\nu} \tag{4}$$

where

$$\varepsilon = \varepsilon^{\alpha}{}_{\alpha} \tag{5}$$

is the trace of the strain tensor obtained by contraction. The volume dilatation ε is defined as the change in volume per original volume [6, see pp. 149–152] and is an invariant of the strain tensor.

It is interesting to note that the structure of Eq.(4) is similar to that of the field equations of General Relativity, viz.

$$R^{\mu\nu} - \frac{1}{2}g^{\mu\nu}R = -KT^{\mu\nu} \tag{6}$$

where $K = 8\pi G/c^4$ and G is the gravitational constant. This strengthens our conjecture that the geometry of the spacetime continuum can be seen as a representation of the deformation of the spacetime continuum resulting from the strains generated by the energy-momentum stress tensor.

Rest-Mass Energy Relation

As shown in [1], the contraction of Eq.(4) yields the relation

$$2(\mu_0 + 2\lambda_0)\varepsilon = T^{\alpha}{}_{\alpha} \equiv T \tag{7}$$

where $T^{\alpha}{}_{\alpha}$ corresponds to the invariant rest-mass energy density

$$T^{\alpha}{}_{\alpha} = T = \rho c^2 \tag{8}$$

where ρ is the rest-mass density. The relation between the invariant volume dilatation ε and the invariant rest-mass energy density is thus given by

$$2(\mu_0 + 2\lambda_0)\varepsilon = \rho c^2 \tag{9}$$

or, in terms of the bulk modulus κ_0,

$$4\kappa_0\varepsilon = \rho c^2. \tag{10}$$

As we noted in [1], this equation demonstrates that rest-mass energy density arises from the volume dilatation of the spacetime continuum. The rest-mass energy is equivalent to the energy required to dilate the volume of the spacetime continuum, and is a measure of the energy stored in the spacetime continuum as volume dilatation. The volume dilatation is an invariant, as is the rest-mass energy density.

Decomposition into Distortions and Dilatations

As also shown in [1], when the strain tensor $\varepsilon^{\mu\nu}$ and the energy-momentum stress tensor $T^{\mu\nu}$ are decomposed into a deviation tensor (the *distortion*) and a scalar (the *dilatation*), the strain-stress relation then becomes separated into dilatation and distortion relations:

$$\begin{aligned} &\text{dilatation} : t = 2(\mu_0 + 2\lambda_0)e = 4\kappa_0 e = \kappa_0\varepsilon \\ &\text{distortion} : t^{\mu\nu} = 2\mu_0 e^{\mu\nu} \end{aligned} \tag{11}$$

where

$$\varepsilon^{\mu\nu} = e^{\mu\nu} + e g^{\mu\nu} \tag{12}$$

with

$$e^{\mu}{}_{\nu} = \varepsilon^{\mu}{}_{\nu} - e\delta^{\mu}{}_{\nu} \tag{13}$$

$$e = \frac{1}{4}\varepsilon^{\alpha}{}_{\alpha} = \frac{1}{4}\varepsilon \tag{14}$$

and similarly

$$T^{\mu\nu} = t^{\mu\nu} + t g^{\mu\nu} \tag{15}$$

with

$$t^{\mu}{}_{\nu} = T^{\mu}{}_{\nu} - t\delta^{\mu}{}_{\nu} \tag{16}$$

$$t = \frac{1}{4}T^{\alpha}{}_{\alpha}. \tag{17}$$

The distortion-dilatation decomposition is evident in the dependence of the dilatation relation on the bulk modulus κ_0 and of the distortion relation on the shear modulus μ_0. The dilatation relation of Eq.(11) corresponds to rest-mass energy, while the distortion relation is traceless and thus massless, and corresponds to shear transverse waves. We also noted in [1] that this decomposition of spacetime continuum deformations into a massive dilatation and a massless transverse wave distortion is somewhat reminiscent of wave-particle duality.

3 Kinematic Relations

The strain $\varepsilon^{\mu\nu}$ can be expressed in terms of the displacement u^{μ} through the kinematic relation [6, see pp. 149–152]:

$$\varepsilon^{\mu\nu} = \frac{1}{2}(u^{\mu;\nu} + u^{\nu;\mu} + u^{\alpha;\mu}u_{\alpha}{}^{;\nu}) \tag{18}$$

where the semicolon (;) denotes covariant differentiation. For small displacements, this expression can be linearized to give the symmetric tensor

$$\varepsilon^{\mu\nu} = \frac{1}{2}(u^{\mu;\nu} + u^{\nu;\mu}) = u^{(\mu;\nu)}. \tag{19}$$

We use the small displacement approximation in this analysis.

An antisymmetric tensor $\omega^{\mu\nu}$ can also be defined from the displacement u^{μ}. This tensor is called the rotation tensor and is defined as [6]:

$$\omega^{\mu\nu} = \frac{1}{2}(u^{\mu;\nu} - u^{\nu;\mu}) = u^{[\mu;\nu]}. \tag{20}$$

Where needed, displacements in expressions derived from Eq.(19) will be written as $u_{\parallel}$ while displacements in expressions derived from Eq.(20) will be written as $u_{\perp}$. Using different symbolic subscripts for these displacements provides a reminder that symmetric displacements are along the direction of motion (longitudinal), while antisymmetric displacements are perpendicular to the direction of motion (transverse).

In general, we have [6]

$$u^{\mu;\nu} = \varepsilon^{\mu\nu} + \omega^{\mu\nu} \tag{21}$$

where the tensor $u^{\mu;\nu}$ is a combination of symmetric and antisymmetric tensors. Lowering index ν and contracting, we get the volume dilatation of the spacetime continuum

$$u^{\mu}{}_{;\mu} = \varepsilon^{\mu}{}_{\mu} = u_{\parallel}{}^{\mu}{}_{;\mu} = \varepsilon \tag{22}$$

where the relation

$$\omega^{\mu}{}_{\mu} = u_{\perp}{}^{\mu}{}_{;\mu} = 0 \tag{23}$$

has been used.

4 Dynamic Equation

4.1 Equilibrium Condition

Under equilibrium conditions, the dynamics of the spacetime continuum is described by the equation [4, see pp. 88–89],

$$T^{\mu\nu}{}_{;\mu} = -X^{\nu} \tag{24}$$

where X^{ν} is the volume (or body) force. As Wald [7, see p. 286] points out, in General Relativity the local energy density of matter as measured by a given observer is well-defined, and the relation

$$T^{\mu\nu}{}_{;\mu} = 0 \tag{25}$$

can be taken as expressing local conservation of the energy-momentum of matter. However, it does not in general lead to a global conservation law. The value $X^{\nu} = 0$ is thus taken to represent the macroscopic local case, while Eq.(24) provides a more general expression.

At the microscopic level, energy is conserved within the limits of the Heisenberg Uncertainty Principle. The volume force may thus be very small, but not exactly zero. It again makes sense to retain the volume force in the equation, and use Eq.(24) in the general case, while Eq.(25) can be used at the macroscopic local level, obtained by setting the volume force X^{ν} equal to zero.

4.2 Displacement Wave Equation

Substituting for $T^{\mu\nu}$ from Eq.(4), Eq.(24) becomes

$$2\mu_0\varepsilon^{\mu\nu}{}_{;\mu} + \lambda_0 g^{\mu\nu}\varepsilon_{;\mu} = -X^{\nu} \tag{26}$$

and, using Eq.(19),

$$\mu_0(u^{\mu;\nu}{}_{\mu} + u^{\nu;\mu}{}_{\mu}) + \lambda_0\varepsilon^{;\nu} = -X^{\nu}. \tag{27}$$

Interchanging the order of differentiation in the first term and using Eq.(22) to express ε in terms of u, this equation simplifies to

$$\mu_0 u^{\nu;\mu}{}_{\mu} + (\mu_0 + \lambda_0)u^{\mu}{}_{;\mu}{}^{\nu} = -X^{\nu} \tag{28}$$

which can also be written as

$$\mu_0\nabla^2 u^{\nu} + (\mu_0 + \lambda_0)\varepsilon^{;\nu} = -X^{\nu}. \tag{29}$$

This is the *displacement wave equation*.

Setting X^{ν} equal to zero, we obtain the macroscopic displacement wave equation

$$\nabla^2 u^{\nu} = -\frac{\mu_0 + \lambda_0}{\mu_0}\varepsilon^{;\nu}. \tag{30}$$

4.3 Continuity Equation

Taking the divergence of Eq.(21), we obtain

$$u^{\mu;\nu}{}_{\mu} = \varepsilon^{\mu\nu}{}_{;\mu} + \omega^{\mu\nu}{}_{;\mu}. \tag{31}$$

Interchanging the order of partial differentiation in the first term, and using Eq.(22) to express u in terms of ε, this equation simplifies to

$$\varepsilon^{\mu\nu}{}_{;\mu} + \omega^{\mu\nu}{}_{;\mu} = \varepsilon^{;\nu}. \tag{32}$$

Hence the divergence of the strain and rotation tensors equals the gradient of the massive volume dilatation, which acts as a source term. This is the continuity equation for deformations of the spacetime continuum.

5 Wave Equations

5.1 Dilatational (Longitudinal) Wave Equation

Taking the divergence of Eq.(28) and interchanging the order of partial differentiation in the first term, we obtain

$$(2\mu_0 + \lambda_0)u^{\mu}{}_{;\mu}{}^{\nu}{}_{\nu} = -X^{\nu}{}_{;\nu}. \tag{33}$$

Using Eq.(22) to express u in terms of ε, this equation simplifies to

$$(2\mu_0 + \lambda_0)\varepsilon^{;\nu}{}_{\nu} = -X^{\nu}{}_{;\nu} \tag{34}$$

or

$$(2\mu_0 + \lambda_0)\nabla^2\varepsilon = -X^{\nu}{}_{;\nu}. \tag{35}$$

Setting X^{ν} equal to zero, we obtain the macroscopic longitudinal wave equation

$$(2\mu_0 + \lambda_0)\nabla^2\varepsilon = 0. \tag{36}$$

The volume dilatation ε satisfies a wave equation known as the dilatational wave equation [6, see p. 260]. The solutions of the homogeneous equation are dilatational waves which are longitudinal waves, propagating along the direction of motion. Dilatations thus propagate in the spacetime continuum as longitudinal waves.

5.2 Rotational (Transverse) Wave Equation

Differentiating Eq.(28) with respect to x^α, we obtain

$$\mu_0 u^{\nu;\mu}{}_{\mu}{}^{\alpha} + (\mu_0 + \lambda_0) u^{\mu}{}_{;\mu}{}^{\nu\alpha} = -X^{\nu;\alpha}. \tag{37}$$

Interchanging the dummy indices ν and α, and subtracting the resulting equation from Eq.(37), we obtain the relation

$$\mu_0 (u^{\nu;\mu}{}_{\mu}{}^{\alpha} - u^{\alpha;\mu}{}_{\mu}{}^{\nu}) = -(X^{\nu;\alpha} - X^{\alpha;\nu}). \tag{38}$$

Interchanging the order of partial differentiations and using the definition of the rotation tensor $\omega^{\nu\alpha}$ of Eq.(20), the following wave equation is obtained:

$$\mu_0 \nabla^2 \omega^{\mu\nu} = -X^{[\mu;\nu]} \tag{39}$$

where $X^{[\mu;\nu]}$ is the antisymmetrical component of the gradient of the volume force defined as

$$X^{[\mu;\nu]} = \frac{1}{2}(X^{\mu;\nu} - X^{\nu;\mu}). \tag{40}$$

Setting X^ν equal to zero, we obtain the macroscopic transverse wave equation

$$\mu_0 \nabla^2 \omega^{\mu\nu} = 0. \tag{41}$$

The rotation tensor $\omega^{\mu\nu}$ satisfies a wave equation known as the rotational wave equation [6, see p. 260]. The solutions of the homogeneous equation are rotational waves which are transverse waves, propagating perpendicular to the direction of motion. Massless waves thus propagate in the spacetime continuum as transverse waves.

5.3 Strain (Symmetric) Wave Equation

A corresponding symmetric wave equation can also be derived for the strain $\varepsilon^{\mu\nu}$. Starting from Eq.(37), interchanging the dummy indices ν and α, adding the resulting equation to Eq.(37), and interchanging the order of partial differentiation, the following wave equation is obtained:

$$\mu_0 \nabla^2 \varepsilon^{\mu\nu} + (\mu_0 + \lambda_0) \varepsilon^{;\mu\nu} = -X^{(\mu;\nu)} \tag{42}$$

where $X^{(\mu;\nu)}$ is the symmetrical component of the gradient of the volume force defined as

$$X^{(\mu;\nu)} = \frac{1}{2}(X^{\mu;\nu} + X^{\nu;\mu}). \tag{43}$$

Setting X^ν equal to zero, we obtain the macroscopic symmetric wave equation

$$\nabla^2 \varepsilon^{\mu\nu} = -\frac{\mu_0 + \lambda_0}{\mu_0} \varepsilon^{;\mu\nu}. \tag{44}$$

This strain wave equation is similar to the displacement wave equation Eq.(30).

6 Discussion and Conclusion

In this paper, we have proposed a framework for the analysis of strained spacetime based on the elastodynamics of the spacetime continuum (*STCED*). In this model, the emphasis is on the displacements of the spacetime continuum infinitesimal elements from their unstrained configuration as a result of the strains applied on the *STC* by the energy-momentum stress tensor, rather than on the geometry of the *STC* due to the energy-momentum stress tensor.

We postulate that this description based on the deformation of the continuum is a description complementary to that of General Relativity which is concerned with modeling the resulting geometry of the spacetime continuum. Interestingly, the structure of the resulting stress-strain relation is similar to that of the field equations of General Relativity. This strengthens our conjecture that the geometry of the spacetime continuum can be seen as a representation of the deformation of the spacetime continuum resulting from the strains generated by the energy-momentum stress tensor. The equivalency of the strain description and of the geometrical description still remains to be demonstrated.

The equilibrium dynamic equation of the spacetime continuum is described by $T^{\mu\nu}{}_{;\mu} = -X^\nu$. In General Relativity, the relation $T^{\mu\nu}{}_{;\mu} = 0$ is taken as expressing local conservation of the energy-momentum of matter. The value $X^\nu = 0$ is thus taken to represent the macroscopic local case, while in the general case, the volume force X^ν is retained in the equation. This dynamic equation leads to a series of wave equations as derived in this paper: the displacement (u^ν), dilatational (ε), rotational ($\omega^{\mu\nu}$) and strain ($\varepsilon^{\mu\nu}$) wave equations.

Hence energy is seen to propagate in the spacetime continuum as deformations of the *STC* that satisfy wave equations of propagation. Deformations can be decomposed into dilatations and distortions. *Dilatations* involve an invariant change in volume of the spacetime continuum which is the source of the associated rest-mass energy density of the deformation. *Distortions* correspond to a change of shape of the spacetime continuum without a change in volume and are thus massless. Dilatations correspond to longitudinal displacements and distortions correspond to transverse displacements of the spacetime continuum.

Hence, every excitation of the spacetime continuum can be decomposed into a transverse and a longitudinal mode of propagation. We have noted that this decomposition into a dilatation with rest-mass energy density and a massless transverse wave distortion, is somewhat reminiscent of wave-particle duality, with the transverse mode corresponding to the wave aspects and the longitudinal mode corresponding to the particle aspects.

A continuity equation for deformations of the spacetime continuum is derived; we find that the divergence of the strain and rotation tensors equals the gradient of the massive volume dilatation, which acts as a source term.

The nature of the spacetime continuum volume force remains to be investigated. In addition, the displacement, dilatational, rotational and strain inhomogeneous wave equations need further investigation.

Submitted on: November 2, 2012 / Accepted on: November 8, 2012

References

1. Millette P.A. On the Decomposition of the Spacetime Metric Tensor and of Tensor Fields in Strained Spacetime. *Progress in Physics*, 2012, v. 4, 5–8.
2. Tartaglia A. A Strained Space-time to Explain the large Scale Properties of the Universe. *International Journal of Modern Physics: Conference Series*, 2011, v. 3, 303–311.
3. Tartaglia A., Radicella N., Sereno M. Lensing in an elastically strained space-time. *Journal of Physics: Conference Series*, 2011, v. 283, 012037.
4. Flügge W. Tensor Analysis and Continuum Mechanics. Springer-Verlag, New York, 1972.
5. Kaku M. Quantum Field Theory; A Modern Introduction. Oxford University Press, Oxford, 1993.
6. Segel L.A. Mathematics Applied to Continuum Mechanics. Dover Publications, New York, 1987.
7. Wald R.M. General Relativity. The University of Chicago Press, Chicago, 1984.

Oblique-Length Contraction Factor in the Special Theory of Relativity

Florentin Smarandache
University of New Mexico 705 Gurley Ave. Gallup, NM 87301, USA. E-mail: smarand@unm.edu

In this paper one generalizes the Lorentz Contraction Factor for the case when the lengths are moving at an oblique angle with respect to the motion direction. One shows that the angles of the moving relativistic objects are distorted.

1 Introduction

According to the Special Theory of Relativity, the Lorentz Contraction Factor is referred to the lengths moving along the motion direction. The lengths which are perpendicular on the direction motion do not contract at all [1].

In this paper one investigates the lengths that are oblique to the motion direction and one finds their Oblique-Length Contraction Factor [3], which is a generalization of the Lorentz Contraction Factor (for $\theta = 0$) and of the perpendicular lengths (for $\theta = \pi/2$). We also calculate the distorted angles of lengths of the moving object.

2 Length-Contraction Factor

Length-Contraction Factor $C(v)$ is just Lorentz Factor:

$$C(v) = \sqrt{1 - \frac{v^2}{c^2}} \in [0, 1] \quad for\ v \in [0, 1] \tag{1}$$

$$L = L' \cdot C(v) \tag{2}$$

where L = non-proper length (length contracted), L' = proper length. $C(0) = 1$, meaning no space contraction [as in Absolute Theory of Relativity (ATR)].

$C(c) = 0$, which means according to the Special Theory of Relativity (STR) that if the rocket moves at speed 'c' then the rocket length and laying down astronaut shrink to zero! This is unrealistic.

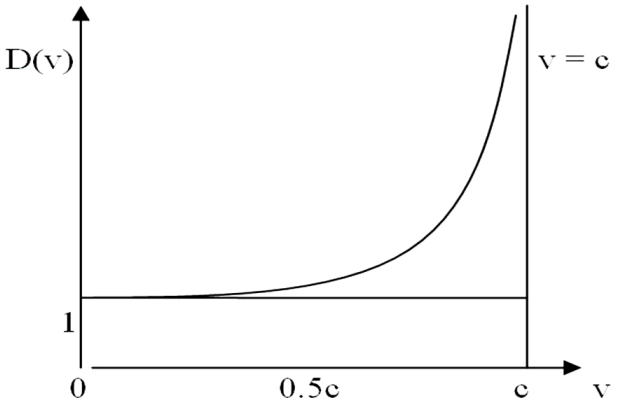

Fig. 1: The graph of the Time-Dilation Factor

3 Time-Dilation Factor

Time-Dilation Factor D(v) is the inverse of Lorentz Factor:

$$D(v) = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \in [1, +\infty] \quad for\ v \in [0, c] \tag{3}$$

$$\Delta t = \Delta t' \cdot D(v) \tag{4}$$

where Δt = non-proper time and, $\Delta t'$ = proper time. $D(0) = 1$, meaning no time dilation [as in Absolute Theory of Relativity (ATR)]; $D(c) = \lim_{v \to c} D(v) = +\infty$, which means according to the Special Theory of Relativity (STR) that if the rocket moves at speed 'c' then the observer on earth measures the elapsed non-proper time as infinite, which is unrealistic. $v = c$ is the equation of the vertical asymptote to the curve of $D(v)$.

4 Oblique-Length Contraction Factor

The Special Theory of Relativity asserts that all lengths in the direction of motion are contracted, while the lengths at right angles to the motion are unaffected. But it didn't say anything about lengths at oblique angle to the motion (i.e. neither perpendicular to, nor along the motion direction), how would they behave? This is a generalization of Galilean Relativity, i.e. we consider the oblique lengths. The length contraction factor in the motion direction is:

$$C(v) = \sqrt{1 - \frac{v^2}{c^2}}. \tag{5}$$

Suppose we have a rectangular object with width W and length L that travels at a constant speed v with respect to an observer on Earth.

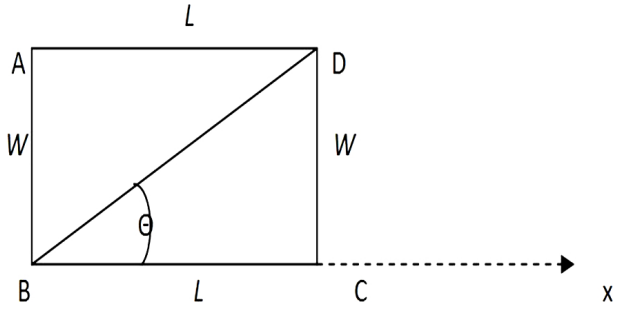

Fig. 2: A rectangular object moving along the x-axis

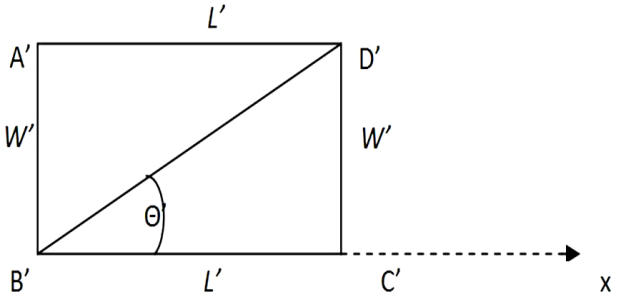

Fig. 3: Contracted lengths of the rectangular object moving along the x-axis

Then its lengths contract and its new dimensions will be L' and W': where $L' = L \cdot C(v)$ and $W' = W$. The initial diagonal of the rectangle ABCD is:

$$\begin{aligned} \delta &= |AC| = |BD| = \sqrt{L^2 + W^2} \\ &= \sqrt{L^2 + L^2 \tan^2\theta} = L\sqrt{1 + \tan^2\theta} \end{aligned} \tag{6}$$

while the contracted diagonal of the rectangle $A'B'C'D'$ is:

$$\begin{aligned} \delta' &= |A'C'| = |B'D'| \\ &= \sqrt{(L')^2 + (W')^2} = \sqrt{L^2 \cdot C(v)^2 + W^2} \\ &= \sqrt{L^2 C(v)^2 + L^2 \tan^2\theta} = L\sqrt{C(v)^2 + \tan^2\theta}. \end{aligned} \tag{7}$$

Therefore the lengths at oblique angle to the motion are contracted with the oblique factor

$$\begin{aligned} OC(v,\theta) &= \frac{\delta'}{\delta} = \frac{L\sqrt{C(v)^2 + \tan^2\theta}}{L\sqrt{1 + \tan^2\theta}} \\ &= \sqrt{\frac{C(v)^2 + \tan^2\theta}{1 + \tan^2\theta}} = \sqrt{C(v)^2\cos^2\theta + \sin^2\theta} \end{aligned} \tag{8}$$

which is different from C(v).

$$\delta' = \delta \cdot OC(v,\theta) \tag{9}$$

where $0 \le OC(v,\theta) \le 1$.

For unchanged constant speed v, the greater is θ in $\left(0, \frac{\pi}{2}\right)$ the larger gets the oblique-length contradiction factor, and reciprocally. By oblique length contraction, the angle

$$\theta \in \left(0, \frac{\pi}{2}\right) \cup \left(\frac{\pi}{2}, \pi\right) \tag{10}$$

is not conserved.

In Fig. 4 the horizontal axis represents the angle θ, while the vertical axis represents the values of the Oblique-Length Contraction Factor $OC(v,\theta)$ for a fixed speed v. Hence $C(v)$ is thus a constant in this graph. The graph, for v fixed, is

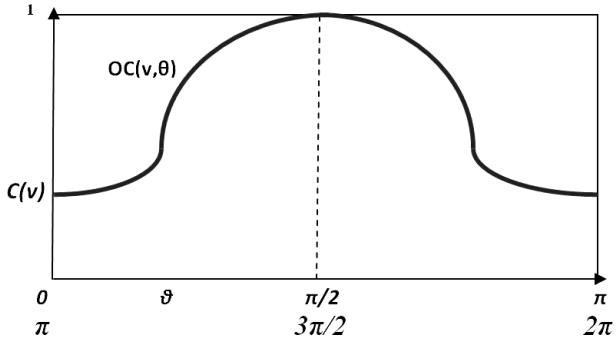

Fig. 4: The graph of the Oblique-Length Contraction Factor $OC(v,\theta)$

periodic of period π, since:

$$\begin{aligned} OC(v,\pi+\theta) &= \sqrt{C(v)^2\cos^2(\pi+\theta) + \sin^2(\pi+\theta)} \\ &= \sqrt{C(v)^2[-\cos\theta]^2 + [-\sin\theta]^2} \\ &= \sqrt{C(v)^2\cos^2\theta + \sin^2\theta} \\ &= OC(v,\theta). \end{aligned} \tag{11}$$

More exactly about the $OC(v,\theta)$ range:

$$OC(v,\theta) \in [C(v), 1] \tag{12}$$

but since $C(v) \in [0,1]$, one has:

$$OC(v,\theta) \in [0,1]. \tag{13}$$

The Oblique-Length Contractor

$$OC(v,\theta) = \sqrt{C(v)^2\cos^2\theta + \sin^2\theta} \tag{14}$$

is a generalization of Lorentz Contractor $C(v)$, because: when $\theta = 0$ or the length is moving along the motion direction, then $OC(v,0) = C(v)$. Similarly

$$OC(v,\pi) = OC(v,2\pi) = C(v). \tag{15}$$

Also, if $\theta = \frac{\pi}{2}$, or the length is perpendicular on the motion direction, then $OC(v,\pi/2) = 1$, i.e. no contraction occurs. Similarly $OC(v, \frac{3\pi}{2}) = 1$.

5 Angle Distortion

Except for the right angles $(\pi/2, 3\pi/2)$ and for the 0, π, and 2π, all other angles are distorted by the Lorentz transform.

Let's consider an object of triangular form moving in the direction of its bottom base (on the x-axis), with speed v, as in Fig. 5:

$$\theta \in \left(0, \frac{\pi}{2}\right) \cup \left(\frac{\pi}{2}, \pi\right) \tag{16}$$

is not conserved.

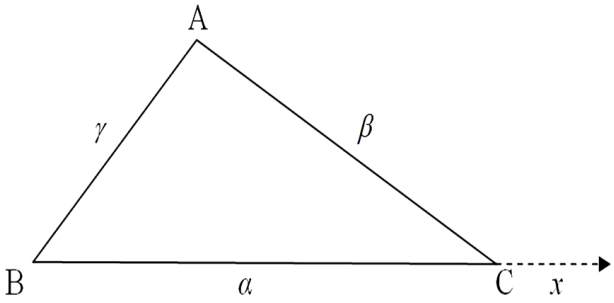

Fig. 5:

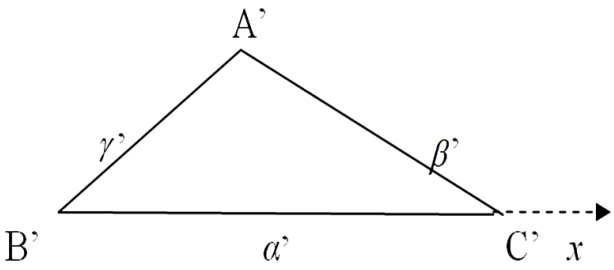

Fig. 6:

The side $|BC| = \alpha$ is contracted with the contraction factor $C(v)$ since BC is moving along the motion direction, therefore $|B'C'| = \alpha \cdot C(v)$. But the oblique sides AB and CA are contracted respectively with the oblique-contraction factors $OC(v, \measuredangle B)$ and $OC(v, \measuredangle \pi - C)$, where $\measuredangle B$ means angle B:

$$\left|A'B'\right| = \gamma \cdot OC(v, \measuredangle B) \tag{17}$$

and

$$\left|C'A'\right| = \beta \cdot OC(v, \measuredangle \pi - C) = \beta \cdot OC(v, \measuredangle A + B) \tag{18}$$

since

$$\measuredangle A + \measuredangle B + \measuredangle C = \pi. \tag{19}$$

Triangle ABC is shrunk and distorted to $A'B'C'$ as in Fig. 6.

Hence one gets:

$$\begin{aligned} \alpha' &= \alpha \cdot C(v) \\ \beta' &= \beta \cdot OC(v, \measuredangle A + B) \\ \gamma' &= \gamma \cdot OC(v, \measuredangle B) \end{aligned} \tag{20}$$

In the resulting triangle $A'B'C'$, since one knows all its side lengths, one applies the Law of Cosine in order to find each angle $\measuredangle A'$, $\measuredangle B'$, and $\measuredangle C'$. Therefore:

$$\measuredangle A' = \arccos \frac{-\alpha^2 \cdot C(v)^2 + \beta^2 \cdot OC(v, \measuredangle A + B)^2 + \gamma^2 \cdot OC(v, \measuredangle B)^2}{2\beta \cdot \gamma \cdot OC(v, \measuredangle B) \cdot OC(v, \measuredangle A + B)}$$

$$\measuredangle B' = \arccos \frac{\alpha^2 \cdot C(v)^2 - \beta^2 \cdot OC(v, \measuredangle A + B)^2 + \gamma^2 \cdot OC(v, \measuredangle B)^2}{2\alpha \cdot \gamma \cdot OC(v) \cdot OC(v, \measuredangle B)}$$

$$\measuredangle C' = \arccos \frac{\alpha^2 \cdot C(v)^2 + \beta^2 \cdot OC(v, \measuredangle A + B)^2 - \gamma^2 \cdot OC(v, \measuredangle B)^2}{2\alpha \cdot \beta \cdot OC(v) \cdot OC(v, \measuredangle A + B)}.$$

As we can see, the angles $\measuredangle A'$, $\measuredangle B'$, and $\measuredangle C'$ are, in general, different from the original angles A, B, and C respectively.

The distortion of an angle is, in general, different from the distortion of another angle.

Submitted on November 12, 2012 / Accepted on November 15, 2012

References

1. Einstein A. On the Electrodynamics of Moving Bodies. *Annalen der Physik*, 1905, v. 17, 891–921.
2. Smarandache F. Absolute Theory of Relativity and Parameterized Special Theory of Relativity and Noninertial Multirelativity. Somipress, Fes, 1982.
3. Smarandache F. New Relativistic Paradoxes and Open Questions. Somipress, Fes, 1983.

Caducity of Idea about Wave Function Collapse as well New Views on Schrödinger's Cat and Quantum Measurements

Spiridon Dumitru

(Retired) Department of Physics, "Transilvania" University, B-dul Eroilor 29, 500036 Brašov, Romania, Phone: +40 746 058 152.
E-mail: s.dumitru42@yahoo.com

Investigated idea was actuated by the old opinion that a measurement of a quantum observable should be regarded a as a single deterministic sampling. But, according to the last decades studies, such observables are veritable random variables and their measurements must imply significant sets of statistical samplings. So one finds the indubitable caducity of the approached idea. Contiguously the respective finding allows to put into a new light the controversial questions like the Schrödinger cat thought experiment or description of quantum measurements.

1 Introduction

A recent highly authorized opinion [1] points out the existing deadlock that: *"There is now ... no entirely satisfactory interpretation of Quantum Mechanics (QM)"*. As major question of that deadlock is recognized as being [2] the problem of Quantum Measurements (QMS), in whose center still stands [3] the Idea about Wave Function Collapse (IWFC). For IWFC, demarcated as above, the most known debates and mainstream publications are reported in [1–3].

Here, in discussing the IWFC question, we try to present a somewhat "unconventional" strategy based on viewpoints promoted in our modest researches about QM, developed over last few decades (see [4, 5] and references).

Firstly we note the fact that, historically, IWFC emerged at the same time with the inaugural ideas regarding the Conventional Interpretation of Uncertainty Relations (CIUR). In the main CIUR started [4, 5] by mixing the theoretical representation (modeling) of a a physical quantity regarding a quantum state/system with a "*fictitious observation*" (done through some thought (gedanken) measuring experiment) of the respective quantity. The mentioned mixing invented and promoted the widespread term of "*observable*" for such a quantity. Below, similarly to the nowadays publications, we will use also the respective term.

After the alluded start CIUR coagulates in a form of an apparent doctrine centered on two main pieces:

(i) Heisenberg's thought-experimental formula and

(ii) Robertson- Schrödinger theoretical relation.

The respective doctrine can be incorporated [4,5] in few basic items (presumptions/ assertions). A deep analysis shows [4, 5] that the respective items, considered as single or grouped pieces, are incriminated by indubitable facts which are unsurmountable within the framework of CIUR. Then CIUR proves oneself to be deprived of necessary qualities for a valid scientific construction. Consequently, in spite of its apology in many modern texts (see references from [4]), CIUR must be abandoned as a wrong conception without any real value or scientific significance.

In its turn, IWFC continued to be present in important publications (see [1–3] and references), with explicit or implicit references to CIUR. It was aroused by the conflict between two items:

(i) The old opinion that a measurement of a quantum observable should be regarded a as a single deterministic sampling and

(ii) The agreement, enforced by theoretical practice, that studies of quantum systems use probabilistic (non-deterministic) entities (wave functions and observables/operators).

For avoiding conflict and breaking a deadlock it was devised the IWFC which, in different readings, was assumed in a large number of publications. But, as a rule, such assumptions were (and still are) not associated with adequate investigations regarding the truthfulness of the respective idea in relation with the QM questions. A modest investigation of that kind we will try to present below in the next sections.

Firstly, in Section 2, we point out the fact that in the main (i.e. irrespectively of its readings) IWFC is nothing but an useless fiction. Such a fact certainly shows the caducity and failure of the respective idea. In Section 3 we discuss the some aspects contiguous between failure of IWFC and famous subject of Schrödinger's cat thought experiment. Then within Section 4 we argue that alternatively to the IWFC we have to reconsider our views about QM theory in relation with QMS. So, for the readings of the respective theory, we must to consider either a restricted-QM (r-QM) or an extended-QM (e-QM) form. On the one hand the r-QM is essentially the version promoted by usual QM textbooks [6, 7] and it deals exclusively only with the modeling of intrinsic properties for the studied systems. On the other hand e-QM must to contain also obligatorily some additional elements regarding QMS descriptions (i.e. theoretical models about characteristics of measuring devices/procedures). Figuratively speaking e-QM consists in r-QM united with QMS descriptions. An simple exemplification of a QMS description, regarded in the mentioned sense, is presented in the end of the same Section 4. Fi-

nally, in Section 5, are given some concluding remarks about the views from this article.

2 Uselessness of IWFC

Now let us try to estimate the usefulness and truthfulness degrees of IWFC. Such an estimation can be obtained if IWFC is regarded through the details of its constituent elements. The before mentioned regard must be opened by observation that the starting purpose of IWFC was to harmonize the following two conflicting Items (***I***):

- ***I***$_1$ The old opinion (of the same time as CIUR) that a measurement of a quantum observable A, specific to a state/system at atomic scale, should be regarded as a single sampling which gives an unique deterministic result, say a_i;
- ***I***$_2$ The theoretical agreement that, due to the probabilistic character of wave function Ψ describing the alluded state/system, the observable A is endowed with a spectrum (set) of distinct values.

So came into an equivocal sight IWFC knew a lot of debates (see [1–3] and references). In essence, the solution promoted by the respective debates can be summarized within the following Subterfuge (***S***):

- ***S*** The unique result a_i and wave function Ψ, mentioned in items ***I***$_1$ and ***I***$_2$, should be seen (and described) through the wave function collapse $\Psi \longmapsto \psi_i$, where Ψ depicts the considered quantum state/system in its wholeness while ψ_i is the a_i-eigenfunction of the operator $\widehat{A}$ (associated to the observable A) — i.e $\widehat{A}\psi_i = a_i\psi_i$.

For a proper judgment of such a subterfuge we have to reconsider the correctness of the items ***I***$_1$ and ***I***$_2$. In the light of such a reason it must to note that studies from the last decades (see [4–7] and references) consolidated beyond doubt the fact that, mathematically, a quantum observable A (through of the operator $\widehat{A}$) is a true random variable. In a theoretical viewpoint, for a given quantum state/system, such a variable is regarded as endowed with a spectra of values associated with corresponding probabilities (more exactly probability amplitudes). Then, from an experimental perspective, a measurement of a quantum observable requires an adequate number of samplings finished through a significant statistical group of data (outcomes).

Previous opinions about the randomness of quantum observables can be consolidated indirectly by mentioning the quantum-classical probabilistic similarity (see [4, 8]) among the respective observables and macroscopic variables studied within phenomenolgical (thermodynamic) theory of fluctuations [4, 9–14]. In this way let us refer to such a macroscopic random observable $\widehat{\mathbb{A}}$. Its intrinsic (*in*) characteristics are given in details by a continuous spectra of values $\mathcal{A}$ inside of spectra (range) Ω_{in} (i.e. $\mathcal{A} \in \Omega_{in}$), associated with a probability density $w_{in} = w_{in}(\mathcal{A})$. Then for $\widehat{\mathbb{A}}$, in its fullness, a single experimental sampling delivering an unique (individual) result, say $\mathcal{A}_i$, is worthlessly. Such a sampling is not described as a collapse of the probability density $w_{in}(\mathcal{A})$. Moreover a true experimental evaluation of $\widehat{\mathbb{A}}$, in its wholeness and regarded equivalently with a stationary random process, requires [15] an adequate lot of samplings finished through a significant statistical set of individual recordings. In a plausible modeling [16, 17] the mentioned recordings (*rec*) can be described by another probability density $w_{rec} = w_{rec}(\mathcal{A})$.

The above notifications about quantum observables point out clearly the complete incorrectness of item ***I***$_1$. Consequently, even if in the main the item ***I***$_2$ is a true assertion, the subterfuge ***S*** supporting IWFC proves oneself to be nothing but an useless recommendation. Additionally note that, in the mainstream of publications (see [1–3] and references), the respective subterfuge is not fortified with thorough (and genuine) descriptions regarding the collapse $\Psi \longmapsto \psi_i$. Evidently that the above revealed facts ***point out the caducity and failure of IWFC***.

The previous discussions about IWFC lead us also to the following more general Remark (***R***)

- ***R*** A random variable should not be assessed (measured) by an unique deterministic sampling (trial) but by a statistical ensemble of samplings.

3 Contiguities with the Schrödinger's cat thought experiment

As it is well known [18] the famous Schrödinger's cat thought experiment is a subject often displayed in debates (more or less scientifically) about the significance/interpretations of QM constituents. The essential element in the respective experiment is represented by a killing single decay of a radioactive atom. But the radioactive decays are random (probabilistic) events. Then the mentioned killing decay is in fact a twin analogue of the single sampling noted above in item ***I***$_1$ in connection with IWFC.

The mentioned analogy motivates us to discuss on some contiguities among questions specific to the alluded experiment and those regarding IWFC. We think that, according to the above remark ***R***, the main point of such motivated discussions is to mark down the following Notification (***N***)

- ***N*** When the variable of interest has random characteristics it is useless (even forbidden) to design experiences or actions that relies solely on a single deterministic sampling of that variable.

In the light of such notification the Schrödinger experiment appears to be noting but just a fiction (figment) without any scientific value. That is why the statements like: "*the Schrödinger cat thought experiment remains a topical touchstone for all interpretations of quantum mechanics*", must be regarded as being worthlessly. (Note that such statements are

present in many science popularization texts, e.g. in the ones disseminated via the internet.)

The above notification ***N***, argued for quantum level, can be also of non-trivial significance (interest) at macroscopic scale. For illustrating such a significance let us refer to the thought experimental situation of a classical (macroscopic) cousin of the Schrödinger cat. The regarded situation can be depicted as follows. The cousin is placed in a sealed box together a flask of poison and an internal macroscopic hammer. The hammer is connected to an macroscopic uncontrollable (unobservable) sensor located within the circular error probable (CEP) of a ballistic projectile trajectory. Note that a ballistic projectile is a missile whose flight is governed by the laws of classical mechanics. CEP is defined as the radius of a circle, centered about the mean, whose boundary is expected to include the landing points of 50% of the launching rounds (for more details about ballistic terminology see [19]). The experiment consists in launching of a single projectile, without any possibility to observe the point where it hits the ground. Also the projectile is equipped with a radio transmitter which signals the flight time. If the sensor is smitten by projectile the hammer is activated releasing the poison that kills the cousin. But as the projectile trajectory has a probabilistic character (mainly due to the external ballistic factors) the hitting point is placed with the probability of 50% within the surface of CEP where the sensor is located. That is why, after the projectile time of flight and without opening the box, one can not know the state of living for the cousin. So the whole situation of the classical cousin is completely analogous with the one of quantum Schrödinger's cat. Therefore the thought experiment with classical cousin makes evident oneself as another fiction without any real significance.

We can add here another circumstance where the above notification ***N*** is taken into account (and put in practice) in a classical context. Namely we think that, in the last analysis, the respective notification is the deep reason of the fact that in practice of the traditional artillery (operating only with ballistic projectiles but not with propelled missiles) for destroying a military objective one uses a considerable (statistical) number of projectiles but not a single one.

4 Contiguities with descriptions of quantum measurements

It is easy to see the fact that the considerations from Section 2 are contiguous with the question of QMS descriptions. Such a fact require directly certain additional comments which we try to present here below. In our opinion the mentioned question must be regarded within a context marked by the following set of Topics (***T***):

$\boldsymbol{T}_1$ In its plenitude the QM theory must be considered in a r-QM respectively in an e-QM reading. Fundamentally, on the one hand, r-QM deals with theoretical models regarding intrinsic properties of quantum (atomically sized) systems. On the other hand e-QM has to take into account both the characteristics of measured observable/system and the peculiarities of measuring devices/procedures;

$\boldsymbol{T}_2$ Within r-QM a situation (state/system) is described completely by its intrinsic (*in*) wave function Ψ_{in} and operators $\widehat{A_k}$ ($k = 1, 2, \ldots, f$), associated to its specific observables A_k. Expression of Ψ_{in} is distinct for each situation while the operators $\widehat{A_k}$ have the same mathematical representation in many situations. The concrete mathematical expression for Ψ_{in} may be obtained either from theoretical studies (e.g. by solving the adequate Schrödinger equation) or from a priori considerations (not supported by factual studies). For a given state/system the observables A_k can be put into sight through a small number of global *in*-descriptors such are: *in*-mean values, *in*-deviations or second or higher order *in*-moments and correlations (for few examples see below);

$\boldsymbol{T}_3$ A true experimental evaluation of quantum observables can be obtained by means of an adequate numbers of samplings finished through significant statistical sets of individual recordings. For an observable the samplings must be done on the same occurrences (i.e. practically on very images of the investigated observable and state/system). As regards a lot of observables a global and easy sight of the mentioned evaluation can be done by computing from the alluded recordings some (experimental) *exp*-quantifiers (of global significance) such are: *exp*-mean, *exp*-deviation respectively *exp*-higher order moments;

$\boldsymbol{T}_4$ Usually, a first confrontation of theory versus experience, is done by comparing side by side the *in*-descriptors and *exp*-quantifiers mentioned above in $\boldsymbol{T}_2$ and $\boldsymbol{T}_3$. Then, if the confrontation is confirmatory, the investigations about the studied observable/system can be noticed as a fulfilled task. If the alluded confirmation does not appear the study may be continued by resorting to one or groups of the following upgradings (***u***):
$\boldsymbol{u}_1$) An amendment for expression of Ψ_{in}, e.g. through solving a more complete Schrödinger equation or using the quantum perturbation theory;
$\boldsymbol{u}_2$) Improvements of experimental devices and procedures;
$\boldsymbol{u}_3$) Addition of a theoretical description for the considered QMS;

$\boldsymbol{T}_5$ Through the extension suggested in above upgrading $\boldsymbol{u}_3$ the study changes its reading from a r-QM into an e-QM vision, in the sense mentioned in topic $\boldsymbol{T}_1$. Such an extension needs to be conceived as a stylized representation through a mathematic modeling so that it to include both intrinsic elements (regarding observables/states/systems) and measuring details. Also if the

upgrading $\boldsymbol{u}_3$ is adopted then a true confrontation of theory versus experience must be done not as it was mentioned in $\boldsymbol{T}_4$ but by putting face to face the predictions of QMS description with the experimental data.

For an illustration of the topics $\boldsymbol{T}_1$–$\boldsymbol{T}_5$ let us regard as a QM system a spin-less quantum particle in a rectilinear and stationary movement along the Ox axis. The QMS problems will be reported to the orbital observables momentum p_x and energy E, denoted generically by A.

In terms of $\boldsymbol{T}_2$ the probabilistic intrinsic (*in*) characteristics of such particle are depicted by orbital wave function $\Psi_{in} = \Psi_{in}(x)$ (where coordinate x covers the range Ω). The observables A are described by the associated operators $\widehat{A}$ according the QM rules [6,7] (i.e. by $\widehat{p}_x = -i\hbar\frac{\partial}{\partial x}$ respectively by the Hamiltonian $\widehat{H}$). Then from the class of global *in*-descriptors regarding such an observable A can be mentioned the *in*-mean-value $\langle A\rangle_{in}$ and *in*- deviation $\sigma_{in}(A)$ defined as follows

$$\left.\begin{aligned} \langle A\rangle_{in} &= \left(\Psi_{in},\ \widehat{A}\,\Psi_{in}\right) \\ \sigma_{in}(A) &= \sqrt{\left(\delta_{in}\widehat{A}\,\Psi_{in},\ \delta_{in}\widehat{A}\,\Psi_{in}\right)} \end{aligned}\right\}, \qquad (1)$$

where (f, g) denotes the scalar product of functions f and g, while $\delta_{in}\widehat{A} = \widehat{A} - \langle A\rangle_{in}$.

An actual experimental measurement of observable A in sense of $\boldsymbol{T}_3$ must be done through a set of statistical samplings. The mentioned set gives for A as recordings a collection of distinct values $\{\alpha_1, \alpha_2, \alpha_3, \ldots, \alpha_r\}$ associated with the empirical probabilities (or relative frequencies) $\{\nu_1, \nu_2, \nu_3, \ldots, \nu_r\}$. Usually, for a lower synthesized sight about the mentioned measurement, as experimental (*exp*) quantifiers are chosen the *exp*-mean $\langle A\rangle_{exp}$ and *exp*-deviation $\sigma_{exp}(A)$ given through the formulas:

$$\left.\begin{aligned} \langle A\rangle_{exp} &= \sum_{j=1}^{r} \nu_j \cdot \alpha_j \\ \sigma_{exp}(A) &= \sqrt{\sum_{j=1}^{r} \nu_j \cdot \left(\alpha_j - \langle A\rangle_{exp}\right)^2} \end{aligned}\right\}. \qquad (2)$$

The above considerations about an experimental QMS must be supplemented with the following Observations ($\boldsymbol{O}$):

$\boldsymbol{O}_1$ Note that due to the inaccuracies of experimental devices some of the recorded values $\{\alpha_1, \alpha_2, \alpha_3, \ldots, \alpha_r\}$ can differ from the eigenvalues $\{a_1, a_2, a_3, \ldots, a_s\}$ of the operator $\widehat{A}$.

$\boldsymbol{O}_2$ A comparison at first sight between theory and experiment can be done by putting side by side the corresponding aggregate (global) entities (1) and (2). When one finds that the values of compared entities are in near equalities, usually is admitted the following couple of linked beliefs ($\boldsymbol{b}$):

$\boldsymbol{b}_1$) Theory is pretty correct and
$\boldsymbol{b}_2$) Measuring devices/procedures are almost ideal.
Thus, practically, the survey of debated QMS can be regarded as a finished task.

$\boldsymbol{O}_3$ If instead of the mentioned equalities one detects (one or two) flagrant differences at least one of the alluded beliefs ($\boldsymbol{b}_1$) and ($\boldsymbol{b}_2$) is deficient (and unsustainable). Such a deadlock can be avoided by one or groups of the upgradings $\boldsymbol{u}_1$–$\boldsymbol{u}_3$ mentioned above within the topic $\boldsymbol{T}_4$.

Generally speaking the the upgradings $\boldsymbol{u}_1$–$\boldsymbol{u}_2$ are appreciated and worked (explicitly or implicitly) in mainstream literature (see [1–3] and references). But note that, as far as know, for $\boldsymbol{u}_3$ such an appreciation was neither taken into account nor developed in details in the respective literature. It is our modest task to present below a brief exemplification of upgrading $\boldsymbol{u}_3$ in relationship with the QMS question. The presentation is done in some simple terms of information transmission theory.

An information theory modeling for QMS description

In a QMS process the input information regarding the intrinsic (*in*) properties of the measured system is converted in predicted (*pd*) or output information incorporated within the data received on a device recorder. That is why a QMS appears as an *information transmission process* in which the measuring device plays the role of a *information transmission channel*. So the QMS considered above can be symbolized as $\Psi_{in} \Rightarrow \Psi_{pd}$ for the wave function while the operator $\widehat{A}$ remains invariant. Such symbolization is motivated by the facts that, on the one hand the wave function Ψ is specific for each considered situation (state/system) whereas, on the other hand the operator $\widehat{A}$ preserves the same mathematical expression in all (or at least in many) situations. Note that the (quantity of) information is connected with probability densities $\rho_\eta(x)$ and currents (fluxes) $j_\eta(x)$ ($\eta = in, pd$) defined in terms of $\Psi_\eta(x)$ as in usual QM [4–7]. Add here the fact that $\rho_\eta(x)$ and $j_\eta(x)$ refer to the positional respectively the motional kinds of probabilities. Experimentally the two kinds of probabilities can be regarded as measurable by distinct devices and procedures. Besides, as in practice, one can suppose that the alluded devices are stationary and linear. Then, similarly with the case of measurements regarding classical random observables [4, 16, 17], in an informational reading, the essence of here discussed QMS description can be compressed [4, 17] through the relations:

$$\left.\begin{aligned} \rho_{pd}(x) &= \int \Gamma(x, x')\,\rho_{in}(x')\,dx' \\ j_{pd}(x) &= \int \Lambda(x, x')\,j_{in}(x)\,dx' \end{aligned}\right\}. \qquad (3)$$

Here the kernels $\Gamma(x, x')$ and $\Lambda(x, x')$ include as noticeable parts some elements about the peculiarities of measuring devices/procedures. Mathematically, $\Gamma(x, x')$ and $\Lambda(x, x')$ are normalized in respect with both x and x'. Note that QMS becomes nearly ideal when both $\Gamma(x, x') \to \delta(x - x')$ and $\Lambda(x, x') \to \delta(x - x')$, ($\delta(x - x')$ being the Dirac's δ function). In all other cases QMS appear as non-ideal.

By means of the probability density $\rho_{pd}(x)$ and current $j_{pd}(x)$ can be computed [4] some useful expressions like $\Psi^*_{pd}(x)\,\widehat{A}\,\Psi_{pd}(x)$. Then, for observable A, it is possible to evaluate global indicators of predicted (pd) nature such are pd-mean $\langle A\rangle_{pd}$ and pd-deviation $\sigma_{pd}(A)$ defined, similarly with (1), as follows

$$\left.\begin{aligned} \langle A\rangle_{pd} &= \left(\Psi_{pd}, \widehat{A}\Psi_{pd}\right) \\ \sigma_{pd}(A) &= \sqrt{\left(\delta_{pd}\widehat{A}\Psi_{pd}, \delta_{pd}\widehat{A}\Psi_{pd}\right)} \end{aligned}\right\}. \tag{4}$$

If as regards a quantum observable A, besides a true experimental evaluation, for its measuring process one resorts to a (theoretical/informational) QMS description of the above kind the pd-indicators (4) must be tested by comparing them with their experimental (factual) correspondents (i.e. exp-quantifiers) given in (2).

When the test is confirmatory both theoretical descriptions, of r-QM intrinsic properties of system respectively of QMS, can be considered as adequate and therefore the scientific task can be accepted as finished. But, if the alluded test is of invalidating type, at least one of the mentioned descriptions must be regarded as inadequate and the whole question requires further investigations.

For an impressive illustration of the above presented informational QMS description we consider as observable of interest the energy $A = E = H$ regarding a QM harmonic oscillator. The operator $\widehat{H}$ associated to the respective observable is the Hamiltonian $\widehat{H} = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + \frac{1}{2}m\,\omega^2 x^2$ (m and ω denote the mass respectively the angular frequency of oscillator). The oscillator is considered to be in its lower energetic level, whose intrinsic state is described by the wave function $\Psi_{in}(x) \propto \exp\left\{-\frac{x^2}{4\sigma^2}\right\}$ (here $\sigma = \sigma_{in}(x) = \sqrt{\frac{\hbar}{2m\omega}}$ denote the in-deviation of coordinate x). Then, because Ψ_{in} is a real function, for the considered state one finds $j_{in} = 0$ — i.e. the probability current is absent.

So for the regarded QMS description in (3) remains of interest only first relation dealing with the change $\rho_{in} \to \rho_{pd}$ of the probability density through the kernel $\Gamma(x, x')$. If the supposed measuring device has high performances $\Gamma(x, x')$ can be taken [4] of Gaussian form i.e. $\Gamma(x, x') \propto \exp\left\{-\frac{(x-x')^2}{2\gamma^2}\right\}$, γ being the error characteristic of the respective device. It can been seen that in the case when $\gamma \to 0$ the kernel $\Gamma(x, x')$ degenerates into the Dirac function $\delta(x - x')$. Then $\rho_{pd} = \rho_{in}$. Such a case corresponds to an ideal measurement. Differently, when $\gamma \neq 0$ one speaks of non-ideal measurements.

In the above modeling of QMS description for the energy $A = E = H$ one obtains [4] the following in respectively pd means and deviations

$$\langle H\rangle_{in} = \frac{\hbar\omega}{2}\,; \quad \sigma_{in}(H) = 0\,, \tag{5}$$

$$\langle H\rangle_{pd} = \frac{\omega\left[\hbar^2 + \left(\hbar + 2m\omega\gamma^2\right)^2\right]}{4\left(\hbar + 2m\omega\gamma^2\right)}\,, \tag{6}$$

$$\sigma_{pd}(H) = \frac{\sqrt{2}m\omega^2\gamma^2\left(\hbar + m\omega\gamma^2\right)}{\left(\hbar + 2m\omega\gamma^2\right)}\,. \tag{7}$$

Relations (5) and (7) show that even if Ψ_{in} has the quality of an eigenfunction for $\widehat{H}$ (as $\sigma_{in}(H) = 0$), due to the measurement Ψ_{pd} is deprived of such a quality (because $\sigma_{pd}(H) \neq 0$).

5 Concluding remarks

We point out, on the one hand, the historical emergence of the IWFC from the conflict between the items $\boldsymbol{I}_1$ and $\boldsymbol{I}_2$ mentioned in Section 2. Then we remind the fact that, on the other hand, the modern studies certify the random characteristics of quantum observables. Therefore a true measurement of such an observable requires a whole set of statistically significant samplings. The respective requirement invalidate indubitably the alluded item $\boldsymbol{I}_1$. So IWFC is proved as a caducous and useless recommendation.

Contiguously the respective proof allows to put into a new light the famous Schrödinger's cat thought experiment. We argue in Section 3 that Schrödinger's experiment is noting but just a fiction without any scientific value. The argumentation relies on the notification that: "When the variable of interest has random characteristics it is useless (even forbidden) to design experiences or actions that relies solely on a single deterministic sampling of that variable". The same notification is useful in appreciating of some non-quantum problems such are a Schrödinger's-type experiment with a classical cat or statistical practices in traditional artillery.

The question of IWFC caducity is contiguous also with the problem of QMS descriptions. That is why in Section 4 we present some brief considerations about the respective problem. Thus we propose that QM theory to be regarded either in a r-QM or in an e-QM reading, as it refers to the studied observables and systems without or with taking into account the QMS descriptions. The proposal is consolidated with simple illustration regarding a spin-less quantum oscillator in a rectiliniar and stationary movement along the Ox axis. Particularly we suggest an approach of QMS descriptions based on information transmission theory.

Of course that other different approaches about QMS descriptions can be imagined. They can be taken into account for extending QM theory towards an e-QM reading, as complete/convincing as possible.

Submitted on: November 15, 2012 / Accepted on: November 18, 2012

References

1. Weinberg S. Collapse of the state vector. *Physical Review A*, 2012, v. 85, 062116; arXiv: 1109.6462.
2. Ghirardi G. Collapse theories. *The Stanford Encyclopedia of Philosophy*, Winter 2011 Edition, Stanford.
3. Omnes R. Decoherence and wave function collapse. *Foundations of Physics*, 2011, v. 41, 1857–1880.
4. Dumitru S. Reconsideration of the uncertainty relations and quantum measurements. *Progress in Physics*, 2008, v. 2, 50–68.
5. Dumitru S. Do the uncertainty relations really have crucial significances for physics? *Progress in Physics*, 2010, v. 4, 25–29.
6. Cohen-Tannoudji C., Diu B., Laloe F. Quantum Mechanics. Vol. I. J. Wiley/Herman, NY–Paris, 1977.
7. Schwabl F. Quantum Mechanics. 2nd. rev. ed., Springer, Berlin, 1995.
8. Dumitru S. The Plank and Boltzmann constants as similar generic indicators of stochasticity: some conceptual implications of quantum-nonquantum analogies. *Physics Essays*, 1993, v. 6, 5–20.
9. Munster A. Statistical Thermodynamics. Vol. I, Springer, Berlin, 1969.
10. Landau L., Lifchitz E. Physique Statistique. Mir, Moscou, 1984.
11. Schwabl F. Statistical Mechanics. Springer, Berlin 2002.
12. Dumitru S. Fluctuations and thermodynamic inequalities. *Physica Scripta*, 1974, v. 10, 101–103.
13. Dumitru S., Boer A. Fluctuations in the presence of fields — phenomenological Gaussian approximation and a class of thermodynamic inequalities. *Physical Review E*, 2001, v. 64, 021108.
14. Boer A., Dumitru S. Higher order correlations in the presence of fields. *Physical Review E*, 2002, v. 66, 046116.
15. Korn G.A., Korn T.M. Mathematical Handbook for Scientists and Engineers. Chap. 19, McGraw Hill, New York 1968; Russian version — Nauka, Moscow, 1977.
16. Dumitru S. Phenomenological theory of recorded fluctuations. *Physics Letters A*, 1974, v. 48, 109–110.
17. Dumitru S., Boer A. On the measurements regarding random observables. *Romanian Journal of Physics*, 2008, v. 53, 1111–1116; http://www.nipne.ro/ rjp/2008_53_9-10.html
18. Schrödinger's cat. http://en.wikipedia.org/wiki/Schrödinger's_cat
19. Category: Ballistics. http://en.wikipedia.org/wiki/Category:Ballistics

Gravitational Field Shielding by Scalar Field and Type II Superconductors

B. J. Zhang*, T. X. Zhang†, P. Guggilia‡, and M. Dohkanian§

*Department of Physics and Astronomy, Vanderbilt University, Nashville, TN 37235. E-mail: bojun.zhang@vanderbilt.edu
*Department of Physics, Alabama A & M University, Normal, Alabama 35762.
†Department of Physics, Alabama A & M University, Normal, Alabama 35762. E-mail: tianxi.zhang@aamu.edu
‡Department of Physics, Alabama A & M University, Normal, Alabama 35762.
§Department of Physics, Alabama A & M University, Normal, Alabama 35762.

The gravitational field shielding by scalar field and type II superconductors are theoretically investigated. In accord with the well-developed five-dimensional fully covariant Kaluza-Klein theory with a scalar field, which unifies the Einsteinian general relativity and Maxwellian electromagnetic theory, the scalar field cannot only polarize the space as shown previously, but also flatten the space as indicated recently. The polarization of space decreases the electromagnetic field by increasing the equivalent vacuum permittivity constant, while the flattening of space decreases the gravitational field by decreasing the equivalent gravitational constant. In other words, the scalar field can be also employed to shield the gravitational field. A strong scalar field significantly shield the gravitational field by largely decreasing the equivalent gravitational constant. According to the theory of gravitational field shielding by scalar field, the weight loss experimentally detected for a sample near a rotating ceramic disk at very low temperature can be explained as the shielding of the Earth gravitational field by the Ginzburg-Landau scalar field, which is produced by the type II superconductors. The significant shielding of gravitational field by scalar field produced by superconductors may lead to a new spaceflight technology in future.

1 Introduction

Gravitation is one of the four fundamental interactions of nature. According to the Newtonian universal law of gravitation, any two objects in the universe attract each other with a force that is directly proportional to the product of their masses and inversely proportional to the square of the distance between them. According to the Einsteinian general theory of relativity, gravitation is directly related to the curvature of spacetime. The Schwarzschild solution of the general relativity for a static spherically symmetric body predicts the perihelion precession of planets, the deflection of distant star light by the Sun, the gravitational redshift of Sun's light, and the time delay of radar echoes, which have been well tested by the measurements [1-4].

To study the shielding of the gravitational field in analogous to the shielding of the electromagnetic field, Majorana [5] in 1920 modified the Newtonian gravitational field of an object with a nonzero extinction coefficient $h \neq 0$ as

$$g = g_N \exp\left[-h \int \rho(r) dr\right], \quad (1)$$

where $g_N \equiv G_0 M/r^2$ is the Newtonian gravitational field with G_0 the gravitational constant, M the mass of the object, and r the radial distance from the object center; ρ is the mass density of the object; h is the extinction coefficient. For a spherical object with a constant mass density and radius R, Eq. (1) after integrated becomes

$$g = g_N \exp\left(-\frac{3hM}{4\pi R^2}\right). \quad (2)$$

Laboratory measurements constrained $h \lesssim 10^{-15}$ m^2/kg [6-7]. Space measurements gave $h \lesssim 10^{-19}$ m^2/kg [8-9]. These measurements indicated that the gravitational field shielding is negligible or undetectable in the case of weak fields.

On the other hand, Kaluza [10] in 1921 proposed a five-dimensional (5D) theory to unify the Einsteinian general relativity and Maxwellian electromagnetic theory. The geometric structure and property of the 5D spacetime were then studied by Klein [11-12]. The early Kaluza-Klein (K-K) theory of unification was further developed with a scalar field [13], which can modify both the electromagnetic and gravitational fields. Some previous studies have shown that the scalar field can reduce the electromagnetic field of a charged object and thus polarize the space around the charged object or shield the electromagnetic field from the charged object [14-15]. It is equivalent to increase the free space permittivity constant. Recently, we has shown, in accord with a 5D fully covariant K-K theory, that the scalar field can also reduce the gravitational field of a body and thus flatten the space around the body or shield the gravitational field from the body [16]. It is equivalent to decrease the gravitational constant in and around the body [17].

The scalar field that was introduced to the cosmology in various models has also been considered as a candidate of dark energy for the acceleration of the universe. As the cosmic expansion, the scalar field of the universe changes over time and became repulsive about many years ago and then overcome the gravitational force to accelerate the expansion of the universe. In addition, we have recently shown that a

massive and compact neutron star can generate a strong scalar field, which can significantly shield or reduce its gravitational field, and thus can be more massive and more compact. The mass-radius relation developed under this type of modified gravity with a scalar field can be consistent with the measurements of neutron stars [18].

In this paper, we will investigate the gravitational field shielding by scalar field and type II superconductors. We suggest that the scalar field generated by the type II superconductors has the same physics and thus addable to the scalar field generated by any other types of matter. According to the five-dimensional fully covariant K-K theory with a scalar field, the scalar field of an object can shield its gravity or decrease the equivalent gravitational constant in or around the object. Therefore, the Ginzburg-Lanadu scalar field [19-20] generated by type II superconductors, if it has a similar physics and thus addable to the scalar field of the Earth, can cause a sample to lose a few percent of its weight or the Earth's gravity as detected by [21]. This study will quantitatively analyze the gravitational field shielding due to the scalar field generated by type II superconductors.

2 Gravitational Shielding by Scalar Field

In the 5D fully covariant K-K theory with a scalar field that has successfully unified the 4D Einsteinian general relativity and Maxwellian electromagnetic theory, the gravitational field of a static spherically symmetric object in the Einstein frame was obtained from the 5D equation of motion of matter as [16, 22]

$$g = \frac{c^2}{2\Phi^2}\left(\frac{d\Phi}{dr} + \Phi\frac{d\nu}{dr}\right)e^{\nu-\lambda}, \tag{3}$$

where the metric and scalar field solutions of the 5D fully covariant K-K theory are given by [23]

$$e^{\nu} = \Psi^2\Phi^{-2}, \tag{4}$$

$$e^{\lambda} = \left(1 - \frac{B^2}{r^2}\right)^2\Psi^{-2}, \tag{5}$$

$$\Phi^2 = -\alpha^2\Psi^4 + (1+\alpha^2)\Psi^{-2}, \tag{6}$$

with

$$\Psi = \left(\frac{r-B}{r+B}\right)^{1/\sqrt{3}}, \tag{7}$$

$$B = \frac{G_0 M}{\sqrt{3(1+\alpha^2)}c^2}, \tag{8}$$

$$\alpha = \frac{Q}{2\sqrt{G_0}M}. \tag{9}$$

Here M and Q are the mass and electric charge of the object.

For a neutral object (i.e., $\alpha = 0$ or $Q = 0$), the gravitational field Eq. (3) obtained from the 5D fully covariant K-K theory with a scalar field can be simplified to [17]

$$g = g_N\left(1 - \frac{B^2}{r^2}\right)^{-3}\Phi^{-7} = \frac{1}{64}\left(\Phi^{\sqrt{3}} + 1\right)^6\Phi^{-7-3\sqrt{3}}, \tag{10}$$

where the scalar field Φ and the critical or singular radius B of the K-K solution are simplified as

$$\Phi = \Psi^{-1}, \quad B = \frac{G_0 M}{\sqrt{3}c^2}. \tag{11}$$

The singular radius B of the K-K solution is a factor of $\sqrt{3}/6$ times smaller than the Schwarzschild radius. Eq. (10) indicates that the gravitational field obtained from the 5D fully covariant K-K theory with a scalar field is influenced by the scalar field Φ. This type of influence can be understood as the gravitational field shielding by scalar field.

In the case of weak fields (i.e., $B \ll r$ or in other words, when the gravitational potential energy of a particle is much smaller than the rest energy of the particle), we can approximately simplify g as

$$g = g_N\left(1 - \frac{14G_0 M}{3c^2 r}\right) = 1 - 7\delta\Phi. \tag{12}$$

Here we have replaced $\Phi = 1 + \delta\Phi$. Comparing the field at the surface of object between Eq. (2) and Eq. (12), we obtain the extinction coefficient as

$$h = \frac{56\pi G_0 R}{9c^2} \sim 1.5\times10^{-26}R, \tag{13}$$

which is about $h \sim 1.5\times10^{-26}$ m^2/kg for an object with radius of one meter and about $h \sim 10^{-19}$ m^2/kg for an object with the size of Earth. It is seen that the gravitational field shielding by scalar field is undetectable in a laboratory experiment since the extinction coefficient is very small for an object with laboratory scale size. For an object with Earth's radius $R \sim 6.4\times10^6$ m, the extinction coefficient is $h \sim 10^{-19}$, the order of the space measurements. This analysis is valid only for the case of weak fields.

The reason for the gravitational field to be shed is the significance of the scalar field, which rapidly increases as the radial distance approaches to the singular radius, i.e., $r \to B$ (Top panel of Figure 1). The gravitational field is inversely proportional to the scalar field with a power of $7 - 3\sqrt{3} \sim 1.8$ if $\Phi \gg 1$ as shown in Eq. (10). By writing Eq. (10) as the Newtonian form of the gravitational field

$$g = \frac{GM}{r^2}, \tag{14}$$

where the G is defined as an equivalent gravitational constant

$$G = G_0\left(1 - \frac{B^2}{r^2}\right)^{-3}\Phi^{-7} = \frac{1}{64}\left(\Phi^{\sqrt{3}} + 1\right)^6\Phi^{-7-3\sqrt{3}}. \tag{15}$$

This suggests that the gravitational field shielding occurs because the strong scalar field significantly varies or decreases the equivalent gravitational constant around the object.

To investigate the gravitational shielding by scalar field in the case of strong fields, we plot in the bottom panel of Figure 1 the gravitational field or constant ratio (g/g_N or G/G_0)

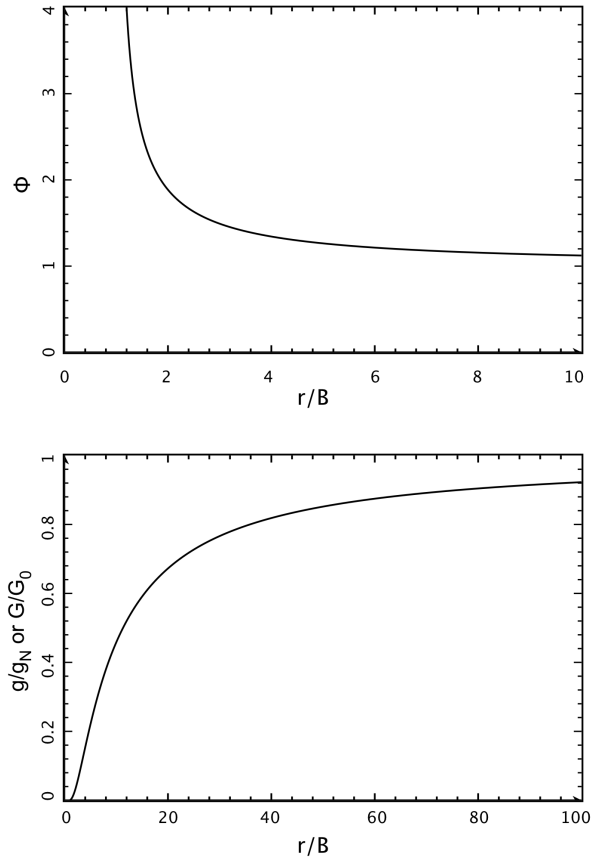

Fig. 1: Scalar field and gravitational field shielding by scalar field. Top panel: the scalar field (Φ) and bottom panel: the ratio between the K-K and Newtonian gravitational field or constant (g/g_N or G/G_0) of a neutral object vs. the normalized radial distance r/B [17].

as a function of the radius distance (r/B) [17]. It is seen that the gravitational field is significantly reduced (or shed) by the scalar field when r is comparable to B. For instances, the gravitational field is shed by $\sim$ 10% (or the percentage of weight loss for a sample object) at $r = 100B$, by $\sim$ 20% at $r = 35B$, by $\sim$ 40% at $r = 15B$, by $\sim$ 80% at $r = 5B$, and $\sim$ 100% at $r = B$. Therefore, for a weak field, the relative difference of the field is small and thus the shielding effect is negligible. For a strong field, however, the gravitational field or constant ratio is small or the relative difference of the field is large so that the shielding effect is significant. The gravitational field of an object, when $r = B$ or its mass-to-radius ratio is about $M/r \simeq 2 \times 10^{27}$ kg/m, is completely shed by the strong scalar field or by the huge amount of mass enclosed. As shown in the top panel of Figure 1, the scalar field increases as r approaches B. The scalar field is ~ 1.4 at $r = 4B$, ~ 4 at $r = 1.6B$, and tends to infinity when $r \to B$. When the scalar field is unity (i.e., $\Phi = 1$), we have $g/g_N = 1$, which refers to that the gravitational field is not shed. When the scalar field significantly departs from the unity, for instance, at $\Phi = 1.2$ or $\delta\Phi = 0.2$, we have $g/g_N \simeq 0.3$, which refers to that a 70% of gravitational field is shed by the scalar field.

3 Gravitational Shielding by Type II Superconductors

About two decades ago, Podkletnov and Nieminen [21] experimentally discovered that a bulk sintered ceramic (type II superconductor) disk of $YBa_2Cu_3O_{7-x}$ can have a moderate shielding effect against the gravitational field. This effect increases with the speed of disk rotation and also depends on the temperature. It was suggested that the shielding effect is the result of a certain state of energy that exists inside the crystal structure of the superconductor at low temperature. This state of energy changes the interactions between electromagnetic, nuclear, and gravitational fields inside a superconductor, and is responsible for the observational phenomena. But a shielding physics has not yet been developed.

Here, we propose a possible shielding physics to explain this phenomena. According to the Ginzburg-Landau theory, a rotating disk of type II superconductor at the phase transition with low temperature (e.g., 70K) generates a scalar field [19-20, 24-30] that varies the equivalent gravitational constant along with the Earth scalar field in and around the superconductor and thus shields the gravitational field of the Earth. According to the 5D fully covariant K-K theory and solution, the scalar field of the Earth at the surface is about the unity because $B \ll r$. Now, in the Podkletnov and Nieminen's experiment, the ceramic (or type II) superconductor can produce an extra scalar field $\delta\Phi$, which is responsible for the small weight loss of the sample.

Based on the previously-developed Landau theory of the second-order phase transition, Ginzburg and Landau [19, 30] showed that the free energy F of a superconductor per unit volume near the transition can be expressed in terms of a complex order parameter field ψ by

$$F = F_n + a|\psi|^2 + \frac{b}{2}|\psi|^4 + \frac{1}{2m}|(-i\hbar \bigtriangledown - 2e\vec{A})\psi|^2 + \frac{|\vec{B}|^2}{2\mu_0}, \quad (16)$$

where F_n is the free energy in the normal phase, a and b are phenomenological parameters, m is an effective mass, e is the charge of electron, $\vec{A}$ is the magnetic vector potential and $\vec{B}$ is the magnetic field. The absolute value of the complex order parameter field $|\psi|$ can be considered as a real scalar field called Ginzburd-Landau scalar field denoted here by $\Phi_{\rm GL} \equiv |\psi|$. Then, in Eq. (16), the second and third terms are the scalar field potential energy; the first part of the fourth term is the scalar field kinetic energy; and the other parts of the fourth term give the energy that couples the scalar field and magnetic field; and the last term is the energy of magnetic field.

By minimizing F with respect to fluctuations of ψ and $\vec{A}$, one can derive the Ginzburg-Landau equations [30-31]

$$a\psi + b|\psi|^2\psi + \frac{1}{2m}(-i\hbar \bigtriangledown - 2e\vec{A})^2\psi = 0, \quad (17)$$

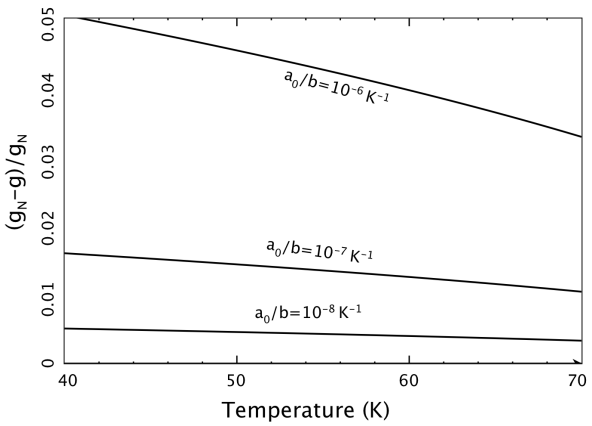

Fig. 2: Gravitational field shielding by scalar field associated with type II superconductor disk. The percentage of weight loss of the sample is plotted as a function of the temperature of the type II superconductors [17].

$$\vec{j} = \frac{2e}{m}\mathrm{Re}[\psi^*(-i\hbar\bigtriangledown - 2e\vec{A})\psi], \tag{18}$$

where $\vec{j}$ is the electrical current density, which is real.

For a homogeneous superconductor, in which $\vec{j} = 0$, Eq. (17) can be simplified to

$$a\psi + b|\psi|^2\psi = 0. \tag{19}$$

The solution $\psi = 0$ is trivial and corresponds to the normal state of the superconductor above the superconducting transition temperature T_c. The non-trivial solution of Eq. (19) determines the Ginzburg-Landau scalar field

$$\Phi_{\mathrm{GL}} \equiv |\psi| = \sqrt{-\frac{a}{b}} = \sqrt{-\frac{a_0}{b}(T - T_c)}. \tag{20}$$

Here, we have assumed the temperature dependence of a to be $a = a_0(T - T_c)$ with positive ratio a_0/b. For the YBCO superconductor, $T_c \sim 93$ K. Suggesting all types of scalar fields to be similar in physics and addable, we obtain the total scalar field in or around a type II superconductor,

$$\Phi_{\mathrm{total}} = \Phi_{\mathrm{Earth}} + \Phi_{\mathrm{GL}}$$
$$= 1 + \frac{2G_0M_{\mathrm{E}}}{3c^2R_{\mathrm{E}}} + \sqrt{-\frac{a_0}{b}(T - T_c)}, \tag{21}$$

where M_{E} and R_{E} are Earth's mass and radius.

To quantitatively study the gravitational field shielding by the Ginzburg-Landau scalar field along with the Earth scalar field, we plot in Figure 2 the weight relative loss of the sample or the gravitational field relative change at the sample as a function of the temperature of the type II superconductor. It is seen that the weight relative loss of the sample or the gravitational field relative change increases as the temperature decreases or as the ratio a_0/b increases. At $T \sim 70$ K and $a_0/b \sim 10^{-8} - 10^{-6}$, the weight relative loss or the gravitational field relative change is $\sim 0.5 - 3\%$, which can be the order of measurements [21].

4 Discussion and Conclusion

For a rotating disk of type II superconductor, the acceleration of inertially moving cooper pairs in the superconductor is equivalent to a gravitational field, which may couple with the Ginzburg-Landau scalar field to produce an extra shielding effect on gravity as shown in [21]. In future study, we will quantitatively analyze the rotation dependence for the gravitational field shielding by the Ginzburg-Landau scalar field of type II superconductors.

As a consequence, we have analytically studied the gravitational field shielding by scalar field and type II superconductors, in accord with the 5D fully covariant K-K theory with a scalar field and the Ginzburg-Landau theory for superconductors. The results have indicated that the gravitational field shielding by the scalar field of a body is very small at an undetectable level if the field is weak. The extinction coefficient derived from the comparison with the Majorana's gravitational field shielding theory is consistent with laboratory and space measurements. In the case of strong fields, however, the gravitational field shielding effect can be significant. This will have important applications in strong-field astrophysics and greatly impact the physics of supernova explosions, the models of neutron stars for their mass-radius relations, and the theory of black hole formations.

Detection of the gravitational field shielding is a challenge to a laboratory experiment, but possible especially when the object becomes a superconductor. A type II superconductor may produce a significant Ginzburg-Landau scalar field at the phase transition and thus may be used to shield gravity as claimed by [21]. The result obtained from this study can be consistent with the measurements. The significant shielding of gravitational field by scalar field produced by superconductors may lead to a new spaceflight technology in future. The gravitational field shielding by type II superconductors still need further experimentally confirmed.

Acknowledgements

This work was supported by the NASA EPSCoR grant (NNX 07AL52A), NSF CISM grant through Boston University, NS F REU program at AAMU, and AAMU Title III programs and by also the National Natural Science Foundation of China (G40890161). The authors thank reviewers and editors for their scientific reviews and editions.

Submitted on: November 19, 2012 / Accepted on November 23, 2012

References

1. Shapiro I. I. Fourth Test of General Relativity. *Physical Review Letters*, 1964, v. 13, 789–791.
2. Shapiro I. I. et al. Fourth Test of General Relativity: New Radar Result. *Physical Reviewer Letters*, 1971, v. 26, 1132–1135.

3. Will C. M. Theory and Experiment in Gravitational Physics, Cambridge Univ. Press, 1993.
4. Will C. M. The confrontation between general relativity and experiment. *Living Reviews in Relativity*, 2006, v. 9, 5–100.
5. Majorana Q. On gravitation: Theoretical and experimental researches. *Philosophical Magazine*, 1920, v. 39, 488–504.
6. Unnikrishnan C. S., Gillies G. T. New limits on the gravitational Majorana screening from the Zurich G experiment. *Physical Review D*, 2000, v. 61, id. 101101.
7. Caputo M. On new limits of the coefficient of gravitation shielding. *Journal of Astrophysics and Astronomy*, 2006, v. 27, 439–441.
8. Eckhardt D. H. Gravitational shielding. *Physical Review D*, 1990, v. 42, 2144–2145.
9. Yang X. S., Wang Q. S. Gravity anomaly during the Mohe total solar eclipse and new constraint on gravitational shielding parameter. *Astrophysics and Space Science*. 2002, v. 282, 245–253.
10. Kaluza T. On the problem of unity in physics. *Sitzungsber Preuss Aklad. Wiss. Berlin Math. Phys.* K1, 1921, 966–972.
11. Klein O. Quantum theory and five-dimensional theory of relativity. *Zeitschrift für Physik*, 1926a, v. 37, 895–906.
12. Klein O. The atomicity of electricity as a quantum theory law. *Nature*, 1926b, v. 118, 516–516.
13. Jordan P. Erweiterung der projektiven relativitastheorie. *Annual Physics* (Leipzig), 1947, v. 1, 219–228.
14. Nodvik J. S. Suppression of singularities by the g^{55} field with mass and classical vacuum polarization in a classical Kaluza-Klein theory. *Physical Review Letters*, 1985, v. 55, L2519–L2522.
15. Dragilev V. M. Vacuum polarization of a scalar field in anisotropic multidimensional cosmology. *Theoretical Mathematics in Physics*, 1990, v. 84, 887–893.
16. Zhang T. X. Gravitational field shielding and supernova explosion. *Astrophysical Journal Letters*, 2010, v. 725, L117–L121.
17. Zhang B. J. Gravitational field shielding and neutron star mass-radius relation. *Research Report* (Alabama A & M University), 2011.
18. Zhang B. J., Zhang, T. X., Guggila, P., Dokhanian, M. Neutron star mass-radius relation with gravitational field shielding by scalar field-Gravitational field shielding by scalar field and neutron star. *Research in Astronomy and Astrophysics*, 2011, Submitted.
19. Ginzburg V. l., Landau L. D. On the theory of superconductivity. *Zh. Eksp. Teor. Fiz.*, 1950, v. 20, 1064–1082.
20. Robertson G. A. Manipulating the vacuum scalar field with superconductor: A search for exotic material. *AIP Conference Proceedings*, 2005, v. 746, 1371–1378.
21. Podkletnov E., Nieminen R. A possibility of gravitational force shielding by bulk $YBa_2Cu_3O_{7-x}$. *Physica C*, 1992, v. 203, 441–444.
22. Zhang T. X. Gravitationless black hole. *Astrophysics and Space Science*, 2011, v. 334, 311–316.
23. Zhang T. X. Electric redshifts and Quasars. *Astrophysical Journal Letters*, 2006, v. 636, L61–L64.
24. Kajantie K., Karjalainen M., Laine M., Peisa J. Masses and phase structure in the Ginzburg-Landau model. *Physical Review B*, 1998, v. 57, 3011–3016.
25. Kiometzis M., Kleinert H., Schakel A. M. J. Critical exponents of the superconducting phase transition. *Physical Review Letters*, 1994, v. 73, 1975–1977.
26. Rosenstein B., Li D. P. Ginzburd-Landau theory of type II superconductors in magnetic field. *Reviews of Modern Physics*, 2010, v. 82, 109–168. bibitem Li N., Noever D., Robertson T., Koczer R., Brantley, W. Static test for a gravitational force coupled to type II YBCO superconductor, *Physica C*, 1997, v. 281, 260–267.
27. Unnikrishan C. S. Does a superconductor shield gravity? *Physica C*, 1996, v. 266, 133–137.
28. Wu N. Gravitational shielding effect in gauge theory of gravity. *Communications in Theoretical Physics*, 2004, v. 41, 567–572.
29. Rosenstein B., Li D. P. Ginzburd-Landau theory of type II superconductors in magnetic field. *Reviews of Modern Physics*, 2010, v. 82, 109–168.
30. Wikipedia, Ginzburg-Landau theory.
31. Cyrot M. Ginzburg-Landau theory for superconductors. *Reports on Progress in Physics*, 1973, v. 36, 103–158.

The Journal on Advanced Studies in Theoretical and Experimental Physics, including Related Themes from Mathematics

PROGRESS IN PHYSICS

A quarterly issue scientific journal, registered with the Library of Congress (DC, USA). This journal is peer reviewed and included in the abstracting and indexing coverage of: Mathematical Reviews and MathSciNet (AMS, USA), DOAJ of Lund University (Sweden), Zentralblatt MATH (Germany), Scientific Commons of the University of St. Gallen (Switzerland), Open-J-Gate (India), Referativnyi Zhurnal VINITI (Russia), etc.

Electronic version of this journal:
http://www.ptep-online.com

Postal Address

Department of Mathematics and Science,
University of New Mexico,
705 Gurley Ave., Gallup, NM 87301, USA

ISSN: 1555-5534 (print)
ISSN: 1555-5615 (online)

Standard Address Number: 297-5092
Printed in the United States of America

APRIL 2013 **VOLUME 2**

CONTENTS

Information for Authors and Subscribers

Progress in Physics has been created for publications on advanced studies in theoretical and experimental physics, including related themes from mathematics and astronomy. All submitted papers should be professional, in good English, containing a brief review of a problem and obtained results.

All submissions should be designed in LaTeX format using *Progress in Physics* template. This template can be downloaded from *Progress in Physics* home page http://www.ptep-online.com. Abstract and the necessary information about author(s) should be included into the papers. To submit a paper, mail the file(s) to the Editor-in-Chief.

All submitted papers should be as brief as possible. We accept brief papers, no larger than 8 typeset journal pages. Short articles are preferable. Large papers can be considered in exceptional cases to the section *Special Reports* intended for such publications in the journal. Letters related to the publications in the journal or to the events among the science community can be applied to the section *Letters to Progress in Physics*.

All that has been accepted for the online issue of *Progress in Physics* is printed in the paper version of the journal. To order printed issues, contact the Editors.

This journal is non-commercial, academic edition. It is printed from private donations. (Look for the current author fee in the online version of the journal.)

Causal Version of Newtonian Theory by Time–Retardation of the Gravitational Field Explains the Flyby Anomalies

Joseph C. Hafele

Retired Physicist; Home Office: 618 S. 24th St., Laramie, WY, USA
E-mail: cahafele@bresnan.net

Classical Newtonian gravitational theory does not satisfy the causality principle because it is based on instantaneous action-at-a-distance. A causal version of Newtonian theory for a large rotating sphere is derived herein by time-retarding the distance between interior circulating point-mass sources and an exterior field-point. The resulting causal theory explains exactly the six flyby anomalies reported by Anderson et al. in 2008. It also explains exactly an anomalous decrease in the Moon's orbital speed. No other known theory has been shown to explain both the flyby anomalies and the lunar orbit anomaly.

1 Introduction

In 2008 Anderson *et al.* reported that anomalous orbital-energy changes have been observed during six spacecraft flybys of the Earth [1]. The reported speed-changes range from a maximum of +13.28 mm/s for the NEAR flyby to a minimum of −4.6 mm/s for the Galileo-II flyby. Anderson *et al.* also found an empirical prediction formula that gives calculated speed-changes that are close to the observed speed-changes. If the speed-change for the empirical prediction formula is designated by δv_{emp}, it can be expressed as follows

$$\delta v_{\text{emp}} = \frac{2v_{\text{eq}}}{c}\, v_{\text{in}}\, (\cos\lambda_{\text{in}} - \cos\lambda_{\text{out}}) = \\ = -\frac{2v_{\text{eq}}}{c}\, v_{\text{in}} \int_{t_{\text{in}}}^{t_{\text{out}}} \sin(\lambda(t))\, \frac{d\lambda}{dt}\, dt\,, \qquad (1)$$

where v_{eq} is the Earth's equatorial rotational surface speed, c is the vacuum speed of light, v_{in} is the initial asymptotic inbound speed, λ_{in} is the asymptotic inbound geocentric latitude, and λ_{out} is the asymptotic outbound geocentric latitude. If t is the observed coordinate time for the spacecraft in its trajectory, then $\lambda_{\text{in}} = \lambda(t_{\text{in}})$ and $\lambda_{\text{out}} = \lambda(t_{\text{out}})$. If $d\lambda/dt = 0$, then $\delta v_{\text{emp}} = 0$. An order of magnitude estimate for the maximum possible value for δv_{emp} is $2(5\times10^2/3\times10^8)\, v_{\text{in}} \sim 30$ mm/s.

The following is a direct quote from the conclusions of an article published in 2009 by M. M. Nieto and J. D. Anderson [2]:

> "Several physicists have proposed explanations of the Earth flyby anomalies. The least revolutionary invokes dark matter bound to Earth. Others include modifications of special relativity, of general relativity, or of the notion of inertia; a light speed anomaly; or anisotropy in the gravitational field —- all of those, of course, deny concepts that have been well tested. And none of them have made comprehensive, precise predictions of Earth flyby effects. For now the anomalous energy changes observed in Earth flybys remain a puzzle. Are they the result of imperfect understandings of conventional physics and experimental systems, or are they the harbingers of exciting new physics?"

It appears that a new and possibly unconventional theory is needed.

The empirical prediction formula found by Anderson *et al.* is not based on any mainstream theory (it was simply "picked out of the air"), but it is remarkably simple and does produce calculated speed-changes that are surprisingly close to the observed speed-changes. The formula for δv_{emp} (1) gives three clues for properties that need to be satisfied by any theory that is developed to explain the flyby anomaly: 1) the theory must produce a speed-change that is proportional to the ratio v_{eq}/c, 2) the anomalous force acting on the spacecraft must change the λ component of the spacecraft's speed, and 3) the speed-change must be proportional to v_{in}.

The objective of this article is threefold: 1) derive a new causal version of classical acausal Newtonian theory, 2) show that this new version is able to produce exact agreement with all six of the anomalous speed-changes reported by Anderson *et al.*, and 3) show that it is also able to explain exactly a "lunar orbit anomaly" that will be described below. The proposed new version for Newtonian theory requires only mainstream physics: 1) classical Newtonian theory and 2) the causality principle which requires time-retardation of the gravitational field. It also satisfies the three requirements of the empirical prediction formula.

The proposed theory is based on a simple correction that converts Newton's acausal theory into a causal theory. The resulting causal theory has a new, previously overlooked, time-retarded transverse component, designated $\boldsymbol{g}_{\text{trt}}$, which depends on $1/c_{\text{g}}$, where c_{g} is the speed of gravity, which approximately equals the speed of light. The new total gravitational field for a large spinning sphere, $\boldsymbol{g}$, has two components, the standard well-known classical acausal radial component, $\boldsymbol{g}_{\text{r}}$, and a new relatively small time-retarded transverse vortex component, $\boldsymbol{g}_{\text{trt}}$. The total vector field $\boldsymbol{g} = \boldsymbol{g}_{\text{r}} + \boldsymbol{g}_{\text{trt}}$. The zero-divergence vortex transverse vector field $\boldsymbol{g}_{\text{trt}}$ is orthogonal to the irrotational radial vector field $\boldsymbol{g}_{\text{r}}$.

The new total vector field is consistent with Helmholtz's theorem, which states that any physical vector field can be expressed as the sum of the gradient of a zero-rotational scalar potential and the curl of a zero-divergence vector potential [3]. This means that $\boldsymbol{g}_{\mathrm{r}}$ can be derived in the standard way from the gradient of a scalar potential, and $\boldsymbol{g}_{\mathrm{trt}}$ can be derived from the curl of a vector potential, but $\boldsymbol{g}_{\mathrm{trt}}$ cannot be derived from the gradient of a scalar potential.

The proposed causal version can be derived by using the slow-speed weak-field approximation for general relativity theory.

2 Summary of the derivation of the formulas for the time-retarded transverse gravitational field and the predicted flyby speed-changes

In the section entitled *The Linear Approximation to GR* in W. Rindler's popular textbook *Essential Relativity* [4], Rindler derives the formulas for the time-retarded scalar potential φ, the time-retarded "gravitoelectric" acceleration field $\boldsymbol{e}$, the time-retarded vector potential $\boldsymbol{a}$, and the time-retarded "gravitomagnetic" induction field $\boldsymbol{h}$. His formulas for φ, $\boldsymbol{e}$, $\boldsymbol{a}$, and $\boldsymbol{h}$ are derived from general relativity theory by using the slow-speed weak-field approximation. They are as follows

$$\left.\begin{aligned} &\varphi = G\iiint\left[\frac{\rho}{r''}\right]dV, \qquad \boldsymbol{a} = \frac{G}{c}\iiint\left[\frac{\rho\,\boldsymbol{u}}{r''}\right]dV \\ &\boldsymbol{e} = -\,\boldsymbol{\nabla}\varphi\,, \qquad\qquad\quad \boldsymbol{h} = \boldsymbol{\nabla}\times 4\,\boldsymbol{a} \end{aligned}\right\}, \quad (2)$$

where ρ is the mass-density of the central object, $\boldsymbol{u}$ is the inertial velocity (the velocity in an inertial frame) of a source-point-mass in the central object, $\boldsymbol{r}''$ is the vector distance from an inner source-point-mass to an outer field-point, and the square brackets [ ] mean that the enclosed function is to be evaluated at the retarded time, i.e., the time retarded by the light travel time from the source-point to the field-point.

Let the origin for an inertial (nonaccelerating and nonrotating) frame-of-reference coincide with the center-of-mass of a contiguous central object. Let $\boldsymbol{r}'$ be the radial vector from the origin to a source-point-mass in the central object, and let $\boldsymbol{r}$ be the radial vector from the origin to an external field-point, so that $\boldsymbol{r}'' = \boldsymbol{r} - \boldsymbol{r}'$. The square brackets in the triple integrals in (2) indicate that the integrands $[\rho/r'']$ and $[\rho\boldsymbol{u}/r'']$ are to be integrated over the volume of the central object at the retarded time.

Let m be the mass of a test-mass that occupies the field-point at $\boldsymbol{r}$, and let $\boldsymbol{v}$ be the inertial velocity of the test mass. The analogous Lorentz force law, i.e., the formula for the time-retarded gravitational force $\boldsymbol{F}$ acting on m at $\boldsymbol{r}$, is [4]

$$\boldsymbol{F} = -\,m\left(\boldsymbol{e} + \frac{1}{c}\,(\boldsymbol{v}\times\boldsymbol{h})\right) = -\,m\,\boldsymbol{\nabla}\left(G\iiint\left[\frac{\rho}{r''}\right]dV\right) - \\ -\,m\left(\boldsymbol{v}\times\left(\boldsymbol{\nabla}\times\left(\frac{4G}{c^2}\iiint\left[\frac{\rho\,\boldsymbol{u}}{r''}\right]dV\right)\right)\right). \quad (3)$$

Rindler's time-retarded version for the slow-speed weak-field approximation gives a complete stand-alone time-retarded solution. The time-retarded fields were derived from general relativity theory, but there is no need for further reference to the concepts and techniques of general relativity theory. Needed concepts and techniques are those of classical Newtonian theory.

Furthermore, Rindler's formulas satisfy the causality principle because the fields are time-retarded. Rindler's version gives a good first approximation only if

$$v^2 \ll c^2, \qquad u^2 \ll c^2, \qquad \frac{GM}{r} = |\varphi| \ll c^2, \quad (4)$$

where M is the total mass of the central object.

Notice in (3) that the acceleration caused by the gravitoelectric field $\boldsymbol{e}$ is independent of c, but the acceleration caused by the gravitomagnetic induction field $\boldsymbol{h}$ is reduced by the factor $1/c^2$. The numerical value for c is on the order of 3×10^8 m/s. If the magnitude for $\boldsymbol{e}$ is on the order of 10 m/s^2 (the Earth's field at the surface), and the magnitudes for $\boldsymbol{u}$ and $\boldsymbol{v}$ are on the order of 10^4 m/s, the relative magnitude for the acceleration caused by $\boldsymbol{h}$ would be on the order of $10\times4\,(10^4/3\times10^8)^2\mathrm{m/s}^2 \sim 10^{-8}\mathrm{m/s}^2$. This estimate shows that, for slow-speed weak-field practical applications in the real world, the acceleration caused by $\boldsymbol{h}$ is totally negligible compared to the acceleration caused by $\boldsymbol{e}$.

The empirical formula indicates that the flyby speed-change is reduced by $1/c$, not by $1/c^2$, which rules out the gravitomagnetic field as a possible cause for the flyby anomalies. The acceleration of $\boldsymbol{h}$ is simply too small to explain the flyby anomalies.

Consequently, the practicable version for Rindler's Lorentz force law becomes the same as a time-retarded version for Newton's well-known inverse-square law

$$\boldsymbol{F} = -\,Gm\,\boldsymbol{\nabla}\iiint\left[\frac{\rho}{r''}\right]dV, \quad (5)$$

where $\boldsymbol{F}$ is the time-retarded gravitational force acting on m.

Let $d^3\boldsymbol{F}$ be the time-retarded elemental force of an elemental point-mass source dm'. The time-retarded version for Newton's inverse-square law becomes

$$d^3\boldsymbol{F} = -\,Gm\,\frac{dm'}{r''^2}\,\frac{\boldsymbol{r}''}{r''}\,, \quad (6)$$

where $\boldsymbol{r}''/r''$ is a unit vector directed towards increasing $\boldsymbol{r}''$.

By definition, the gravitational field of a source at $\boldsymbol{r}'$ is the gravitational force of the source dm' that acts on a test-mass m at $\boldsymbol{r}$ per unit mass of the test-mass. The traditional symbol for the Newtonian gravitational field is $\boldsymbol{g}$. Therefore, the formula for the time-retarded elemental gravitational field $d^3\boldsymbol{g}$ of an elemental point-mass-source at $\boldsymbol{r}'$ for a field-point at $\boldsymbol{r}$ becomes

$$d^3\boldsymbol{g} = \frac{d^3\boldsymbol{F}}{m} = -\,G\,\frac{dm'}{r''^2}\,\frac{\boldsymbol{r}''}{r''}\,. \quad (7)$$

The negative sign indicates that the gravitational force is attractive.

Let t be the observed coordinate time at $\boldsymbol{r}$, let t' be the retarded time at $\boldsymbol{r}'$, and let $c_{\rm g}$ be the speed of propagation of the gravitational field. The connection between t and t' is

$$t = t' + \frac{r''}{c_{\rm g}}\,, \quad t' = t - \frac{r''}{c_{\rm g}}\,. \tag{8}$$

Obviously, $t(t')$ is a function of t', and *vice versa*, $t'(t)$ is a function of t. The Jacobian for the transformation from t to t' is given by

$$\text{Jacobian} = \frac{dt}{dt'} = 1 + \frac{1}{c_{\rm g}}\frac{dr''}{dt'}\,. \tag{9}$$

This Jacobian is needed to solve the triple integral over the volume of the central object. It leads to the necessary factor $1/c_{\rm g}$, where $c_{\rm g}$ is the speed of propagation of the Earth's gravitational field [5].

Let $\rho(\boldsymbol{r}')$ be the mass-density of the central object at $\boldsymbol{r}'$. Then

$$dm' = \rho(\boldsymbol{r}')\,dV. \tag{10}$$

The resulting formula for the elemental total gravitational field $d^3\boldsymbol{g}$, which consists of the radial component $d^3\boldsymbol{g}_{\rm r}$ and the transverse component $d^3\boldsymbol{g}_{\rm trt}$, becomes $d^3\boldsymbol{g} = d^3\boldsymbol{g}_{\rm r} + d^3\boldsymbol{g}_{\rm trt}$. The differential formulas for each component become

$$d^3\boldsymbol{g}_{\rm r} = -G\,\frac{dm'}{r''^2}\left(\frac{\boldsymbol{r}''}{r''}\right)_{\rm r}\,, \qquad d^3\boldsymbol{g}_{\rm trt} = -G\,\frac{dm'}{r''^2}\left(\frac{\boldsymbol{r}''}{r''}\right)_{\rm trt}\,, \tag{11}$$

where $(\boldsymbol{r}''/r'')_{\rm r}$ is the radial component of the unit vector and $(\boldsymbol{r}''/r'')_{\rm trt}$ is the transverse component of the unit vector. The total field is obtained by a triple integration over the volume of the central object at the retarded time.

Let (X, Y, Z) be the rectangular coordinates for the inertial frame-of-reference, and let the Z-axis coincide with the spin axis of the central object. Let RC be the relative radial component, and let TC_Z be the magnitude for the Z-axis component of the relative transverse component. As can be seen in Fig. 1, the formulas for RC and TC_Z are related to $\boldsymbol{r}$, $\boldsymbol{r}'$, and $\boldsymbol{r}''$ by

$$\left.\begin{aligned} RC &= \frac{\boldsymbol{r}\cdot\boldsymbol{r}''}{r''r} = \frac{\boldsymbol{r}\cdot\boldsymbol{r} - \boldsymbol{r}\cdot\boldsymbol{r}'}{r''r} \\ TC_Z &= \frac{(\boldsymbol{r}\times\boldsymbol{r}'')_Z}{r''r} = \frac{(\boldsymbol{r}'\times\boldsymbol{r})_Z}{r''r} = \frac{r'_X r_Y - r'_Y r_X}{r''r} \end{aligned}\right\}, \tag{12}$$

where r_X, r_Y are the X, Y components of $\boldsymbol{r}$, and r'_X, r'_Y are the X, Y components of $\boldsymbol{r}'$.

The formula for the magnitude of $\boldsymbol{g}_{\rm trt}$ becomes [5]

$$\boldsymbol{g}_{\rm trt} = \iiint\left(-G\,\frac{dm'}{r''^2}\right)(TC_Z)\,(\text{Jacobian})\,. \tag{13}$$

The triple integral is rather easy to solve by using numerical integration if the central object can be approximated by a large spinning isotropic sphere. To get a good first approximation, the Earth was simulated in [5] by a large spinning isotropic sphere.

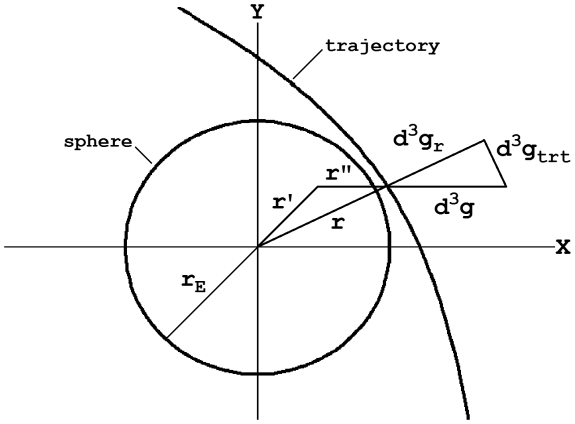

Fig. 1: Depiction of the vector distances $\boldsymbol{r}$, $\boldsymbol{r}'$, and $\boldsymbol{r}''$ and the components of the vector field $d^3\boldsymbol{g}$, $d^3\boldsymbol{g}_{\rm r}$, and $d^3\boldsymbol{g}_{\rm trt}$.

The formulas for the geocentric radial distance to the field-point and its derivative are

$$\left.\begin{aligned} r(\theta) &= \frac{r_{\rm p}\,(1+\varepsilon)}{1+\varepsilon\cos\theta} \\ \frac{dr}{d\theta} &= \frac{r(\theta)^2}{r_{\rm p}}\,\frac{\varepsilon}{1+\varepsilon}\,\sin\theta \end{aligned}\right\}, \tag{14}$$

where θ is the parametric polar coordinate angle for the spacecraft in the plane of the trajectory, $r_{\rm p}$ is the geocentric radial distance at perigee, and ε is the eccentricity of the trajectory.

It is shown in [5] that the formula for the Jacobian is

$$\begin{aligned} \text{Jacobian} &= 1 + \frac{1}{c_{\rm g}}\frac{dr''}{dt} = \\ &= 1 - \frac{r}{c_{\rm g}}\,\frac{r'}{r''}\left(\Omega_\phi - \Omega_{\rm E}\right)\cos\lambda'\,\sin\phi'. \end{aligned} \tag{15}$$

It is also shown in [5] that the only term for $d^3\boldsymbol{g}_{\rm trt}$ that will survive the triple integration is

$$d^3 g_{\rm trt} = -G\bar{\rho}\,r_{\rm E}\,\frac{\Omega_\phi - \Omega_{\rm E}}{\Omega_{\rm E}}\cos^2(\lambda)\,IG\,\frac{dr'}{r_{\rm E}}\,d\lambda'\,d\phi', \tag{16}$$

where $\bar{\rho}$ is the mean value for $\rho(r')$ and the formula for the integrand is

$$IG = \frac{r_{\rm E}^3}{r^3}\,\frac{\rho(r')}{\bar{\rho}}\,\frac{r'^4}{r_{\rm E}^4}\cos^3(\lambda')\frac{\sin^2\phi'}{(1+x)^2} \tag{17}$$

where the variable x is defined by

$$x \equiv \frac{r'^2}{r^2} - 2\,\frac{r'}{r}\cos\lambda'\,\cos\phi'. \tag{18}$$

It has been shown in [5] that the resulting formula for the magnitude of the transverse component is

$$g_{\rm trt}(\theta) = -G\,\frac{I_{\rm E}}{r_{\rm E}^4}\,\frac{v_{\rm eq}}{c_{\rm g}}\,\frac{\Omega_\phi(\theta) - \Omega_{\rm E}}{\Omega_{\rm E}}\,\cos^2(\lambda(\theta))\,PS\,(r(\theta))\,, \tag{19}$$

where G is the gravity constant, I_{E} is the Earth's spherical moment of inertia, r_{E} is the Earth's spherical radius, Ω_{E} is the Earth's spin angular speed, v_{eq} is the Earth's equatorial surface speed, c_{g} is the speed of propagation of the Earth's gravitational field, θ is the spacecraft's parametric polar coordinate angle in the plane of the orbit or trajectory, $\Omega_\theta = d\theta/dt$ is the spacecraft's angular speed, Ω_ϕ is the azimuthal ϕ-component of Ω_θ, λ is the spacecraft's geocentric latitude, r is the spacecraft's geocentric radial distance, and $PS(r)$ is an inverse-cube power series representation for the triple integral over the Earth's volume. The formula for $PS(r)$ is [5]

$$PS(r) \equiv \left(\frac{r_{\mathrm{E}}}{r}\right)^3 \left(C_0 + C_2\left(\frac{r_{\mathrm{E}}}{r}\right)^2 + C_4\left(\frac{r_{\mathrm{E}}}{r}\right)^4 + C_6\left(\frac{r_{\mathrm{E}}}{r}\right)^6\right),$$

where the values for the coefficients are

$$C_0 = 0.50889, \quad C_2 = 0.13931,$$
$$C_4 = 0.01013, \quad C_6 = 0.14671.$$

If the magnitude is negative, i.e., if $\Omega_\phi > \Omega_{\mathrm{E}}$ (prograde), the vector field component $\boldsymbol{g}_{\mathrm{trt}}$ is directed towards the east. If $\Omega_\phi < 0$ (retrograde), it is directed towards the west.

The formula for the time-retarded transverse gravitational field, g_{trt}, satisfies the first requirement of the empirical prediction formula. It is proportional to $v_{\mathrm{eq}}/c_{\mathrm{g}} \cong v_{\mathrm{eq}}/c$. But the empirical prediction formula also requires that the speed-change must be in the λ-component of the spacecraft's velocity, $\boldsymbol{v}_\lambda$. The magnitude for the λ-component is defined by

$$v_\lambda = r_\lambda \frac{d\lambda}{dt} = r_\lambda \frac{d\lambda}{d\theta}\frac{d\theta}{dt} = r_\lambda \Omega_\theta \frac{d\lambda}{d\theta}, \tag{20}$$

where r_λ is the λ-component of r.

The velocity component, $\boldsymbol{v}_\lambda$, is orthogonal to $\boldsymbol{g}_{\mathrm{trt}}$. Consequently, $\boldsymbol{g}_{\mathrm{trt}}$ cannot directly change the magnitude of $\boldsymbol{v}_\lambda$ (it changes the direction).

However, a hypothesized induction-like field, designated $\boldsymbol{F}_\lambda$, can be directed perpendicularly to $\boldsymbol{g}_{\mathrm{trt}}$ in the $\boldsymbol{v}_\lambda$-direction. Assume that the ϕ-component of the curl of $\boldsymbol{F}_\lambda$ equals $-k\, d\boldsymbol{g}_{\mathrm{trt}}/dt$, where k is a constant. This induction-like field can cause a small change in the spacecraft's speed. The reciprocal of the constant k, $v_{\mathrm{k}} = 1/k$, called the "induction speed", becomes an adjustable parameter for each case. The average for all cases gives an overall constant for the causal theory.

The formula for the magnitude of $\boldsymbol{F}_\lambda$ has been shown in [5] to be

$$F_\lambda = \frac{v_{\mathrm{eq}}}{v_{\mathrm{k}}}\frac{r_{\mathrm{E}}}{r(\theta)}\int_0^\theta \frac{r(\theta)}{r_{\mathrm{E}}}\frac{\Omega_\theta(\theta)}{\Omega_{\mathrm{E}}}\frac{1}{r_{\mathrm{E}}}\frac{dr}{d\theta}\frac{dg_{\mathrm{trt}}}{d\theta}\, d\theta. \tag{21}$$

The acceleration caused by $\boldsymbol{F}_\lambda$ satisfies the second requirement of the empirical prediction formula, the one that requires the anomalous force to change the λ-component of the spacecraft's velocity.

The anomalous time rate of change in the spacecraft's orbital energy is given by the dot product, $\boldsymbol{v}\cdot\boldsymbol{F}_\lambda$. It has been shown in [5] that the calculated asymptotic speed-change, δv_{trt}, is given by

$$\delta v_{\mathrm{trt}} = \delta v_{\mathrm{in}} + \delta v_{\mathrm{out}}, \tag{22}$$

where

$$\delta v_{\mathrm{in}} = \delta v(\theta_{\mathrm{min}}), \qquad \delta v_{\mathrm{out}} = \delta v(\theta_{\mathrm{max}}), \tag{23}$$

and

$$\delta v(\theta) = \frac{v_{\mathrm{in}}}{2}\int_0^\theta \frac{r_\lambda(\theta)\, F_\lambda(\theta)}{v_{\mathrm{in}}^2}\frac{d\lambda}{d\theta}\, d\theta. \tag{24}$$

The angles θ_{min} and θ_{max} are the minimum and maximum values for θ. The initial speed $v_{\mathrm{in}} = v(\theta_{\mathrm{min}})$. The speed-change $\delta v(\theta)$ is proportional to v_{in}, which satisfies the third requirement of the empirical prediction formula.

3 Summary of the change in the Moon's orbital speed caused by the Earth's time-retarded transverse gravitational field

In 1995, F. R. Stephenson and L. V. Morrison published a remarkable study of records of eclipses from 700 BC to 1990 AD [6]. They conclude*: 1) the LOD has been increasing on average during the past 2700 years at the rate of $+1.70 \pm 0.05$ ms cy^{-1} (i.e. $(+17.0 \pm 0.5) \times 10^{-6}$ s per year), 2) tidal braking causes an increase in the LOD of $+2.3 \pm 0.1$ ms cy^{-1} (i.e. $(+23 \pm 1)\times 10^{-6}$ s per year), and 3) there is a non-tidal decrease in the LOD, numerically -0.6 ± 0.1 ms cy^{-1} (i.e. $(-6 \pm 1) \times 10^{-6}$ s per year).

Stephenson and Morrison state that the non-tidal decrease in the LOD probably is caused by post-glacial rebound. Post-glacial rebound decreases the Earth's moment of inertia, which increases the Earth's spin angular speed, and thereby decreases the LOD. But post-glacial rebound cannot change the Moon's orbital angular momentum.

According to Stephenson and Morrison, tidal braking causes an increase in the LOD of $(23 \pm 1) \times 10^{-6}$ seconds per year, which causes a decrease in the Earth's spin angular momentum, and by conservation of angular momentum causes an increase in the Moon's orbital angular momentum. It has been shown in [5] that tidal braking alone would cause an increase in the Moon's orbital speed of $(19 \pm 1) \times 10^{-9}$ m/s per year, which corresponds to an increase in the radius of the Moon's orbit of (14 ± 1) mm per year.

But lunar-laser-ranging experiments have shown that the radius of the Moon's orbit is actually increasing at the rate of (38 ± 1) mm per year [7]. This rate for increase in the radius corresponds to an increase in the orbital speed of $(52 \pm 2) \times 10^{-9}$ m/s per year. Clearly there is an unexplained or anomalous difference in the change in the radius of the orbit of (-24 ± 2) mm per year $(38 - 14 = 24)$, and a corresponding anomalous difference in the change in the orbital speed of

*LOD means length-of-solar-day and ms cy^{-1} means milliseconds per century.

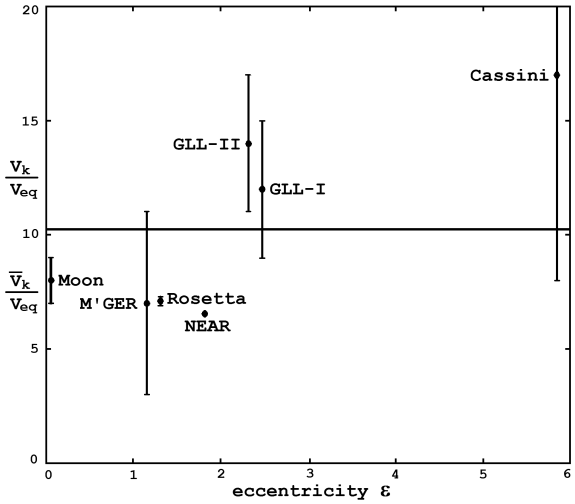

Fig. 2: Required induction speed ratio (designated by •), $v_k/v_{eq} \pm$ an estimate for the uncertainty, versus eccentricity ε. The mean value for all seven ratios is shown by the horizontal line.

$(-33 \pm 3) \times 10^{-9}$ m/s per year $(52 - 19 = 33)$. This "lunar orbit anomaly" cannot be caused by post-glacial rebound, but it can be caused by the proposed causal version of Newton's theory.

It has been shown in [5] that the causal version of Newton's theory produces a change in the Moon's orbital speed of $(-33 \pm 3) \times 10^{-9}$ m/s per year if the value for the induction speed $v_k = (8 \pm 1)\, v_{eq}$. Therefore, the causal theory explains exactly the lunar orbit anomaly. It gives a new closed orbit case for anomalous speed-changes.

4 Calculated speed-changes caused by the time-retarded version of Newton's theory

It has been shown in [5] that the causal version of Newton's theory explains exactly the six flyby anomalies reported by Anderson *et al.* [1]. The required values for v_k cluster between $6\, v_{eq}$ and $17\, v_{eq}$.

A graph of the required induction speed ratios, v_k/v_{eq}, versus eccentricity ε, Fig. 2, shows that the required value for v_k for the lunar orbit anomaly is consistent with the required values for v_k for the six Earth flyby anomalies. The average ± standard deviation is

$$\overline{v}_k = (10 \pm 4)\, v_{eq} = 5 \pm 2 \text{ km/s}. \qquad (25)$$

It will be interesting to compare this average, 5 ± 2 km/s, with parameter values for other flyby theories.

5 Predicted speed-changes for future high-precision Doppler-shift experiments

The speed-change caused by the causal version of Newton's theory depends on the speed of gravity c_g, the properties of the central sphere; mass M_E, radius r_E, angular speed Ω_E, moment of inertia I_E, and equatorial surface speed v_{eq}, on the orbital properties of the spacecraft; radius at perigee r_p, eccentricity ε, inclination to the equatorial plane α_{eq}, and latitude at perigee λ_p, and the induction speed v_k. If $\varepsilon = 0$ or if $\alpha_{eq} = 0$, the speed-change $\delta v_{trt} = 0$. Even if $\varepsilon \neq 0$ and $\alpha_{eq} \neq 0$, δv_{trt} is still equal to zero if perigee is over the equator $(\lambda_p = 0°)$ or one of the poles $(\lambda_p = \pm 90°)$. The maximum speed-change occurs for spacecrafts in highly eccentric and inclined near-Earth orbits.

r_p/r_E	P (hours)	δv_{yr} (mm/s per year)	$\delta v r_{yr}$ (mm/s per year)
2	11.2	+315	−517
3	20.7	+29.5	−76.8
4	31.8	+3.93	−21.0
5	44.4	+0.173	−7.97
6	58.4	−0.422	−3.69
7	73.6	−0.442	−1.95
8	89.9	−0.362	−1.14

Table 1: Calculated period P (in hours) and predicted speed-change for prograde orbits δv_{yr} (in mm/s per year), and the predicted speed-change for retrograde orbits $\delta v r_{yr}$ (in mm/s per year), for a spacecraft in a near-Earth orbit with $\varepsilon = 0.5$, $\alpha_{eq} = 45°$, $\lambda_p = 45°$, $v_k = 14\, v_{eq}$, and for r_p ranging from $2\, r_E$ to $8\, r_E$ [5].

Assume $c_g = c$ and the induction speed is its largest probable value, $v_k = 14\, v_{eq}$. Suppose the orbital properties for a spacecraft are $\varepsilon = 0.5$, $\alpha_{eq} = 45°$, and $\lambda_p = 45°$. Let r_p range from $2\, r_E$ to $8\, r_E$. The period P is given by Kepler's 3rd law, and the annual speed change for prograde $\delta v_{yr} = N_{rev}\, \delta v_{trt}$, and for retrograde $\delta v r_{yr} = N_{rev}\, \delta v_{trt}$, where N_{rev} is the number of revolutions per year. Calculated speed-changes are listed in Table 1 [5].

6 Other theories which explain the Earth flyby anomalies

There are at least two other published theories that explain the Earth flyby anomalies: 1) the 3-space flow theory of R. T. Cahill [8], and 2) the exponential radial field theory of H. J. Busack [9].

In [8] Cahill reviews numerous Michelson interferometer and one-way light-speed experiments which clearly show an anisotropy in the velocity of light. His calculated flyby speed-changes depend on the direction and magnitude for 3-space inflow at the spacecraft on the date and time of the flyby. Cahill found that the average speed for 3-space inflow is 12±5 km/s. Cahill's average, 12±5 km/s, essentially equals the average value for v_k (25), 5±2 km/s.

In [9] Busack applies a small exponential correction for the Earth's radial gravitational field. If $f(\boldsymbol{r}, \boldsymbol{v})$ is Busack's correction, the inverse-square law becomes

$$\boldsymbol{g}_r(\boldsymbol{r}, \boldsymbol{v}) = -\frac{GM_E}{r^2}\frac{\boldsymbol{r}}{r}(1 + f(\boldsymbol{r}, \boldsymbol{v}))\,,$$

where $f(\boldsymbol{r},\boldsymbol{v})$ is expressed as

$$f(\boldsymbol{r},\boldsymbol{v}) = A\exp\left(-\frac{r-r_{\mathrm{E}}}{B-C(\boldsymbol{r}\cdot\boldsymbol{v})/(\boldsymbol{r}\cdot\boldsymbol{v}_{\mathrm{Sun}})}\right).$$

The velocity $\boldsymbol{v}$ is the velocity of the field-point in the "gravitational rest frame in the cosmic microwave background", and $\boldsymbol{v}_{\mathrm{Sun}}$ is the Sun's velocity in the gravitational rest frame. Numerical values for the adjustable constants are approximately $A = 2.2\times10^{-4}$, $B = 2.9\times10^{5}$ m, and $C = 2.3\times10^{5}$ m. Busack found that these values produce rather good agreement with the observed values for the flyby speed-changes.

Both of these alternative theories require a preferred frame-of-reference. Neither has been tested for the lunar orbit anomaly, and neither satisfies the causality principle because neither depends on the speed of gravity.

7 Conclusions and recommendations

This article shows conclusively that the proposed causal version of Newton's theory agrees with the now-known facts-of-observation. It applies only for slow-speeds and weak-fields. Effects of time retardation appear at the relatively large first-order v/c_{g} level, but they have not been seen in the past because they decrease inversely with the cube of the closest distance. If perigee is very close, however, time retardation effects can be relatively large. It is recommended that various available methods be used to detect new observations of effects of the causality principle.

Acknowledgements

I thank Patrick L. Ivers for reviewing many manuscripts and suggesting improvements. I also thank Dr. Robert A. Nelson for bringing to my attention the study of eclipses by F. R. Stevenson and L. V. Morrison [6]. I am especially thankful to Dr. Dmitri Rabounski for encouragement and support for this research project.

A complete version for this research project, containing all necessary details and comprehensive proofs, has been accepted for publication in *The Abraham Zelmanov Journal* [5].

Submitted on: December 07, 2012 / Accepted on: December 08, 2012

References

1. Anderson J.D., Campbell J.K., Ekelund J.E., Ellis J., and Jordon J.F. Anomalous orbital-energy changes observed during spacecraft flybys of Earth. *Physical Review Letters*, 2008, v.100, 091102.
2. Nieto M.M. and Anderson J.D. Earth flyby anomalies. arXiv: gr-qc/0910.1321.
3. Morse P.M. and Feshbach H. Methods of Theoretical Physics. McGraw-Hill, New York, 1953.
4. Rindler W. Essential Relativity, Special, General, and Cosmological. Springer, New York, 1977.
5. Hafele J.C. Earth flyby anomalies explained by time-retarded causal version of Newtonian gravitational theory. *The Abraham Zelmanov Journal*, 2012, v.5 (*under press*).
6. Stephenson F.R. and Morrison L.V. Long-term fluctuations in the Earth's rotation: 700 BC to AD 1990. *Philosophical Transactions of the Royal Society of London A*, 1995, v.351, 165–202.
7. Measuring the Moon's Distance. *LPI Bulletin*, 1994, http://eclipse.gsfc.nasa.gov/SEhelp/ApolloLaser.html
8. Cahill R.T. Combining NASA/JPL one-way optical-fibre light-speed data with spacecraft Earth-flyby Doppler-shift data to characterise 3-space flow. *Progress in Physics*, 2009, v.4, 50–64.
9. Busack H.J. Simulation of the flyby anomaly by means of an empirical asymmetric gravitational field with definite spatial orientation. arXiv: gen-ph/0711.2781.

Dynamical 3-Space Gravitational Waves: Reverberation Effects

Reginald T. Cahill and Samuel T. Deane

School of Chemical and Physical Sciences, Flinders University, Adelaide 5001, Australia. E-mail: Reg.Cahill@flinders.edu.au

Gravity theory missed a key dynamical process that became apparent only when expressed in terms of a velocity field, instead of the Newtonian gravitational acceleration field. This dynamical process involves an additional self-interaction of the dynamical 3-space, and experimental data reveals that its strength is set by the fine structure constant, implying a fundamental link between gravity and quantum theory. The dynamical 3-space has been directly detected in numerous light-speed anisotropy experiments. Quantum matter has been shown to exhibit an acceleration caused by the time-dependence and inhomogeneity of the 3-space flow, giving the first derivation of gravity from a deeper theory, as a quantum wave refraction effect. EM radiation is also refracted in a similar manner. The anisotropy experiments have all shown 3-space wave/turbulence effects, with the latest revealing the fractal structure of 3-space. Here we report the prediction of a new effect, namely a reverberation effect, when the gravitational waves propagate in the 3-space inflow of a large mass. This effect arises from the non-linear dynamics of 3-space. These reverberations could offer an explanation for the Shnoll effect, in which cosmological factors influence stochastic processes, such as radioactive decay rates.

1 Introduction

Newton's inverse square law of gravity, when expressed in terms of an acceleration field $\mathbf{g}(\mathbf{r}, t)$, has the differential form:

$$\nabla \cdot \mathbf{g} = -4\pi G\rho, \quad \nabla \times \mathbf{g} = \mathbf{0} \tag{1}$$

where G is the gravitational constant and ρ is the real matter density. The $\mathbf{g}$ field was believed to exist within an actual Euclidean space. It has become increasingly evident through the observation of spiral galaxies, the expanding universe and gravitational anomalies, that Newton's inverse square law is an incomplete theory of gravity. However a unique generalisation of (1) has lead to a resolution of these anomalies, by writing the acceleration field $\mathbf{g}(\mathbf{r}, t)$ in terms of the Euler acceleration of a velocity field $\mathbf{v}(\mathbf{r}, t)$ [1, 2]:

$$\mathbf{g} = \frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \cdot \nabla)\mathbf{v}, \tag{2}$$

$$\nabla \cdot \left(\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \cdot \nabla)\mathbf{v} \right) = -4\pi G\rho, \quad \nabla \times \mathbf{v} = \mathbf{0}. \tag{3}$$

This approach utilises the the well known Galilean covariant Euler acceleration for a fluidic flow of the substratum with velocity $\mathbf{v}(\mathbf{r}, t)$. The velocity field is defined relative to an observer. The time dependent nature of the flow means that Newtonian gravity, within this flow formalism, can support wave phenomena. But a unique term can be added to (3) that generalises the flow equation, but also preserves the Keplerian nature of the planetary motions that underlie Newton's gravity formalisation:

$$\begin{gathered} \nabla \cdot \left(\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \cdot \nabla)\mathbf{v} \right) + \frac{\alpha}{8} \left((trD)^2 - tr(D^2) \right) = -4\pi G\rho, \\ \nabla \times \mathbf{v} = \mathbf{0}, \quad D_{ij} = \frac{1}{2} \left(\frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i} \right). \end{gathered} \tag{4}$$

Analysis of Bore Hole $\mathbf{g}$ anomaly data revealed that α is the fine structure constant [1]. The additional dynamics explains the "dark matter" effects, and so may be referred to as the dark matter term:

$$\rho_{DM}(\mathbf{r}) = \frac{\alpha}{32\pi G} \left((trD)^2 - tr(D^2) \right) \tag{5}$$

whereby

$$\nabla \cdot \mathbf{g} = \nabla \cdot \left(\frac{\partial \mathbf{v}}{\partial t} + \nabla \left(\frac{\mathbf{v}^2}{2} \right) \right) = -4\pi G \left(\rho_M + \rho_{DM} \right) \tag{6}$$

Dynamical 3-Space is unlike the dualistic space and aether theories of the past, as herein only space exists, and there is no aether. This dynamical and structured space provides an observable and observed local frame of reference. The flow of the dynamical 3-space has been detected many times dating back to the Michelson and Morley 1887 experiment, which has always, until 2002, been mistakenly reported as a null experiment. Wave effects, essentially gravitational waves, are apparent in the data from various anisotropy experiments*. A large part of understanding gravitational waves lies in how they originate, and also in understanding how they propagate. This work herein investigated the propagation of these gravitational waves within the background in-flow of a large mass, such as the earth or the sun. In doing so it was discovered that the dynamics of the propagation resulted in a reverberation effect, caused by the non-linear nature of the flow dynamics, apparent in (3) and (4).

*Vacuum mode Michelson interferometers have zero sensitivity to these waves. So such devices have a fundamental design flaw when intended to detect such waves.

2 3-Space flow dynamics

First we establish the in-flow of space into a spherical mass, assuming for simplicity that the mass is asymptotically at rest, which means that the in-flow has spherical symmetry. In the case of the earth we know that the earth has a large velocity with respect to to the local 3-space frame of reference, some 486 km/s in the direction RA = 4.3^h, Dec = -75° [3]. Here we restrict the analysis to the case of a spherically symmetric inflow into a spherical mass, with density $\rho(r)$ and total mass M. Then (4) becomes ($v' \equiv \partial v(r,t)/\partial r$)

$$\frac{\partial v'}{\partial t} + vv'' + \frac{2vv'}{r} + (v')^2 + \frac{\alpha}{2}\left(\frac{v^2}{2r^2} + \frac{vv'}{r}\right) = -4\pi G\rho \qquad (7)$$

which for a static flow has the exact solution

$$v(r)^2 = \frac{2\beta}{r^{\frac{\alpha}{2}}} + \frac{2G}{(1-\frac{\alpha}{2})r}\int_0^r 4\pi s^2\rho(s)ds + \\ + \frac{2G}{(1-\frac{\alpha}{2})r^{\frac{\alpha}{2}}}\int_r^\infty 4\pi s^{1+\frac{\alpha}{2}}\rho(s)ds, \qquad (8)$$

Here M is the total matter mass, and β is a free parameter. The term $2\beta/r^{\alpha/2}$ describes an inflow singularity or "black hole" with arbitrary strength. This is unrelated to the putative black holes of General Relativity. This corresponds to a primordial black hole. As well the last term in (8) also has a $1/r^{\alpha/2}$ inflow-singularity, but whose strength is mandated by the matter density, and is absent when $\rho(r) = 0$ everywhere. This is a minimal "black hole", and is present in all matter systems. The $2\beta/r^{\alpha/2}$ term will produce a novel long range gravitational acceleration $g = \alpha\beta/2r^{1+\alpha/2}$, as observed in spiral galaxies. For the region outside the sun Keplerian orbits are known to well describe the motion of the planets within the solar system, apart from some small corrections, such as the Precession of the Perihelion of Mercury, which follow from relativistic effects in the more general form of (2) [1]. The case $\beta = 0$ applies to the sun and earth, having only induced "Minimal Attractor" black holes. These minimal black holes contribute to the external $g(r) = GM^*/r^2$ gravitational acceleration, through an effective mass

$$M^* \approx M + \frac{\alpha}{2}M \qquad (9)$$

Outside of a spherical mass, with only an induced black hole, (8) has a solution $v \propto 1/\sqrt{r}$, and then $\rho_{DM} = 0$ outside of the sphere, which explains why the $\alpha-$term in (4) went undiscovered until 2005.

3 Gravitational wave reverberations

We now demonstrate that gravitational waves incoming on, say, a star or planet develop reverberations, in which the wave generates following copies of itself. For numerical accuracy in solving for time dependent effects in (4), we assume a spherically symmetric incoming wave, which is clearly unrealistic, and so find numerical solutions to (7), by using the ansatz $v(r,t) = v(r) + w(r,t)$, where $v(r) \sim -1/\sqrt{r}$ is the static in-flow from (8), applicable outside of the star/planet, and so ignoring the galactic background flow, and where $w(r,t)$ is the wave effect, with the initial wave $w(r,0)$ having the form of a pulse, as shown in fig.1, where the time evolution of $w(r,t)$ is also shown. We see that the initial pulse develops following copies of itself. This is a direct consequence of the non-linearity of (4), or even (3).

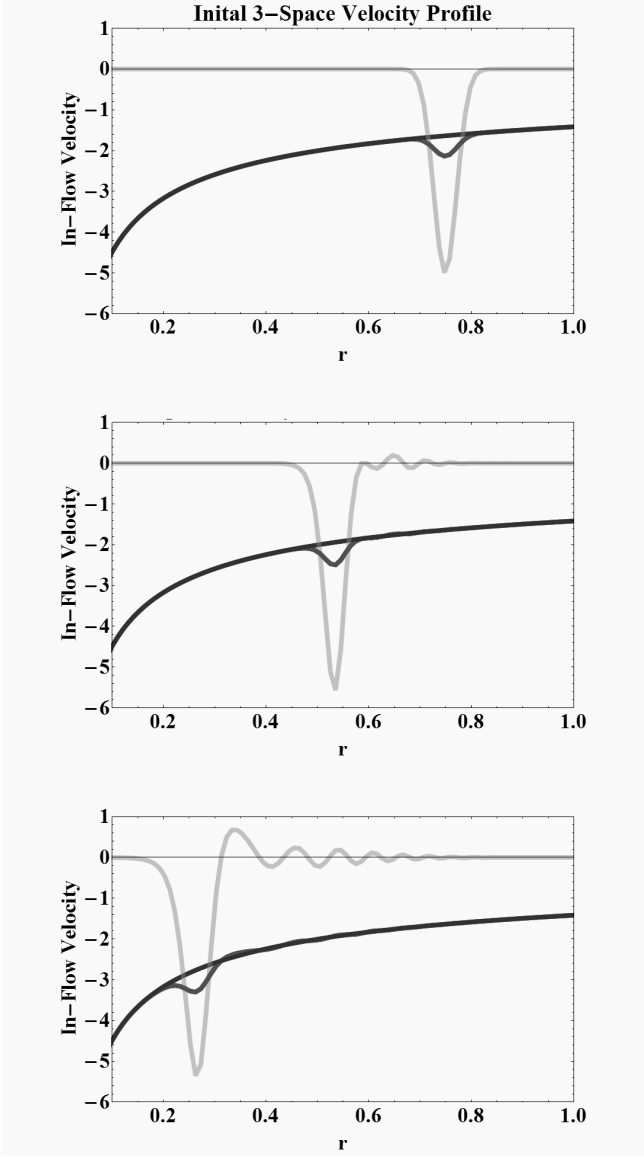

Fig. 1: Inflowing 3-space perturbation $w(r,t)$ (red) and un-perturbed inflow $v(r)$ (blue) velocity profiles outside a mass, with the waveform $w(r,t)$ also shown magnified (yellow).

These reverberations are expected to be detectable in EM speed anisotropy experiments. However because the 3-space

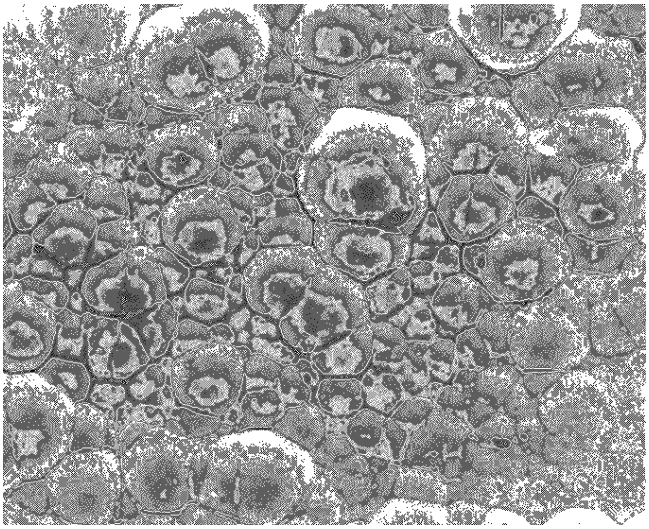

Fig. 2: Representation of the wave data revealing the fractal textured structure of the 3-space, with cells of space having slightly different velocities, and continually changing, and moving with respect to the earth with a speed of ∼500km/s, from [4].

is fractal, as illustrated in fig.2 [4], the reverberations are expected to be complex. As well all systems would generate reverberations, from planets, moons, sun and the galaxy. The timescale for such reverberations would vary considerably. As well as being directly observable in EM anisotropy and gravitational wave detectors, these reverberations would affect, for example, nuclear decay rates, as the magnitude of the 3-space fractal structure is modulated by the reverberations, and this fractal structure will stimulate nuclear processes. Patterns in the decay rates of nuclei have been observed by Shnoll *et al.*, with periodicities over many time scales [5].

The 3-space is detectable because the speed of EM waves, in vacuum is $c \approx 300,000$km/s with respect to that space itself, whereas an observer, in general, will observe anisotropy when the observer is in motion with respect to the space. This effect has been repeatedly observed for over 120 years. The anisotropy detections have always revealed wave/turbulence effects, including the original Michelson-Morley experiment. These wave effects are known as "gravitational waves", although a more appropriate description would be "space waves". In the limit $\alpha \to 0$, (4) and also (7) still have space wave effects, but these generate gravitational wave effects, namely fluctuations in the matter acceleration field $\mathbf{g}(r, t)$, only when $\alpha \neq 0$. So the α-dynamical term is not only responsible for the earth bore hole $\mathbf{g}$ anomaly, and for the so-called "dark matter" effects in spiral galaxies, but can also result in forces acting on matter resulting from the space wave phenomena, and will be large when significant wave effects occur, with large wave effects being essentially a galactic effect.

4 Acknowledgements

This report is from the Flinders University Gravitational Wave Detector Project, funded by an Australian Research Council Discovery Grant: *Development and Study of a New Theory of Gravity*. Special thanks to Dr David Brotherton-Ratcliffe and Professor Igor Bray for ongoing support.

Submitted on: December 7, 2012 / Accepted on December 10, 2012

References

1. Cahill R.T. Process Physics: From Information Theory to Quantum Space and Matter. Nova Science Pub., New York, 2005.
2. Cahill R.T. Dynamical 3-Space: A Review. *Ether Space-time and Cosmology: New Insights into a Key Physical Medium, Duffy M. and Lévy J., eds.*, Apeiron, Montreal, 2009, 135–200.
3. Cahill R.T. Combining NASA/JPL One-Way Optical-Fiber Light-Speed Data with Spacecraft Earth-Flyby Doppler-Shift Data to Characterise 3-Space Flow, *Progress in Physics*, 2009, v. 4, 50–64.
4. Cahill R.T. Characterisation of Low Frequency Gravitational Waves from Dual RF Coaxial-Cable Detector: Fractal Textured Dynamical 3-Space. *Progress in Physics*, 2012, v. 3, 3–10.
5. Shnoll S.E. Cosmophysical Factors in Stochastic Processes. American Research Press, Rehoboth, New Mexico, USA, 2012.

Derivation of Electromagnetism from the Elastodynamics of the Spacetime Continuum

Pierre A. Millette

University of Ottawa (alumnus), Ottawa, Canada. E-mail: PierreAMillette@alumni.uottawa.ca

We derive Electromagnetism from the Elastodynamics of the Spacetime Continuum based on the identification of the theory's antisymmetric rotation tensor with the electromagnetic field-strength tensor. The theory provides a physical explanation of the electromagnetic potential, which arises from transverse (shearing) displacements of the spacetime continuum, in contrast to mass which arises from longitudinal (dilatational) displacements. In addition, the theory provides a physical explanation of the current density four-vector, as the 4-gradient of the volume dilatation of the spacetime continuum. The Lorentz condition is obtained directly from the theory. In addition, we obtain a generalization of Electromagnetism for the situation where a volume force is present, in the general non-macroscopic case. Maxwell's equations are found to remain unchanged, but the current density has an additional term proportional to the volume force.

1 Introduction

Since Einstein first published his Theory of General Relativity in 1915, the problem of the unification of Gravitation and Electromagnetism has been and remains the subject of continuing investigation (see for example [1–9] for recent attempts). The Elastodynamics of the Spacetime Continuum [10, 11] is based on the application of a continuum mechanical approach to the spacetime continuum *(STC)*. Electromagnetism is found to come out naturally from the theory in a straightforward manner.

In this paper, we derive Electromagnetism from the Elastodynamics of the Spacetime Continuum (*STCED*). This theory thus provides a unified description of the spacetime deformation processes underlying general relativistic Gravitation [11] and Electromagnetism, in terms of spacetime continuum displacements resulting from the strains generated by the energy-momentum stress tensor.

1.1 A note on units and constants

In General Relativity and in Quantum Electrodynamics, it is customary to use "geometrized units" and "natural units" respectively, where the principal constants are set equal to 1. The use of these units facilitates calculations since cumbersome constants do not need to be carried throughout derivations. In this paper, all constants are retained in the derivations, to provide insight into the nature of the equations being developed.

In addition, we use rationalized MKSA units for Electromagnetism, as the traditionally used Gaussian units are gradually being replaced by rationalized MKSA units in more recent textbooks (see for example [12]). Note that the electromagnetic permittivity of free space ϵ_{em}, and the electromagnetic permeability of free space μ_{em} are written with "*em*" subscripts as the "0" subscripts are used in *STCED* constants. This allows us to differentiate between for example μ_{em}, the electromagnetic permeability of free space, and μ_0, the Lamé elastic constant for the shear modulus of the spacetime continuum.

2 Theory of Electromagnetism from *STCED*

2.1 Electromagnetic field strength

In the Elastodynamics of the Spacetime Continuum, the antisymmetric rotation tensor $\omega^{\mu\nu}$ is given by [11]

$$\omega^{\mu\nu} = \frac{1}{2}\left(u^{\mu;\nu} - u^{\nu;\mu}\right) \tag{1}$$

where u^μ is the displacement of an infinitesimal element of the spacetime continuum from its unstrained position x^μ. This tensor has the same structure as the electromagnetic field-strength tensor $F^{\mu\nu}$ defined as [13, see p. 550]:

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu \tag{2}$$

where A^μ is the electromagnetic potential four-vector $(\phi, \vec{A})$, ϕ is the scalar potential and $\vec{A}$ the vector potential.

Identifying the rotation tensor $\omega^{\mu\nu}$ with the electromagnetic field-strength tensor according to

$$F^{\mu\nu} = \varphi_0 \omega^{\mu\nu} \tag{3}$$

leads to the relation

$$A^\mu = -\frac{1}{2}\varphi_0 u^\mu_\perp \tag{4}$$

where the symbolic subscript $\perp$ of the displacement u^μ indicates that the relation holds for a transverse displacement (perpendicular to the direction of motion) [11].

Due to the difference in the definition of $\omega^{\mu\nu}$ and $F^{\mu\nu}$ with respect to their indices, a negative sign is introduced, and is attributed to (4). This relation provides a physical explanation

for the electromagnetic potential: it arises from transverse (shearing) displacements of the spacetime continuum, in contrast to mass which arises from longitudinal (dilatational) displacements of the spacetime continuum [11]. Sheared spacetime is manifested as electromagnetic potentials and fields.

2.2 Maxwell's equations and the current density four-vector

Taking the divergence of the rotation tensor of (1), gives

$$\omega^{\mu\nu}{}_{;\mu} = \frac{1}{2}\,(u^{\mu;\nu}{}_{\mu} - u^{\nu;\mu}{}_{\mu}). \tag{5}$$

Recalling (28) from Millette [11], viz.

$$\mu_0 u^{\nu;\mu}{}_{\mu} + (\mu_0 + \lambda_0) u^{\mu}{}_{;\mu}{}^{\nu} = -X^{\nu} \tag{6}$$

where X^{ν} is the volume force and λ_0 and μ_0 are the Lamé elastic constants of the spacetime continuum, substituting for $u^{\nu;\mu}{}_{\mu}$ from (6) into (5), interchanging the order of partial differentiation in $u^{\mu;\nu}{}_{\mu}$ in (5), and using the relation $u^{\mu}{}_{;\mu} = \varepsilon^{\mu}{}_{\mu} = \varepsilon$ from (19) of [11], we obtain

$$\omega^{\mu\nu}{}_{;\mu} = \frac{2\mu_0 + \lambda_0}{2\mu_0}\,\varepsilon^{;\nu} + \frac{1}{2\mu_0}\,X^{\nu}. \tag{7}$$

As seen in [11], in the macroscopic local case, the volume force X^{ν} is set equal to zero to obtain the macroscopic relation

$$\omega^{\mu\nu}{}_{;\mu} = \frac{2\mu_0 + \lambda_0}{2\mu_0}\,\varepsilon^{;\nu} \tag{8}$$

Using (3) and comparing with the covariant form of Maxwell's equations [14, see pp. 42–43]

$$F^{\mu\nu}{}_{;\mu} = \mu_{em} j^{\nu} \tag{9}$$

where j^{ν} is the current density four-vector $(c\varrho, \vec{j})$, ϱ is the charge density scalar, and $\vec{j}$ is the current density vector, we obtain the relation

$$j^{\nu} = \frac{\varphi_0}{\mu_{em}}\,\frac{2\mu_0 + \lambda_0}{2\mu_0}\,\varepsilon^{;\nu}. \tag{10}$$

This relation provides a physical explanation of the current density four-vector: it arises from the 4-gradient of the volume dilatation of the spacetime continuum. A corollary of this relation is that massless (transverse) waves cannot carry an electric charge or produce a current.

Substituting for j^{ν} from (10) in the relation [15, see p. 94]

$$j^{\nu} j_{\nu} = \varrho^2 c^2, \tag{11}$$

we obtain the expression for the charge density

$$\varrho = \frac{1}{2}\,\frac{\varphi_0}{\mu_{em} c}\,\frac{2\mu_0 + \lambda_0}{2\mu_0}\,\sqrt{\varepsilon^{;\nu}\varepsilon_{;\nu}} \tag{12}$$

or, using the relation $c = 1/\sqrt{\epsilon_{em}\mu_{em}}$,

$$\varrho = \frac{1}{2}\,\varphi_0 \epsilon_{em} c\,\frac{2\mu_0 + \lambda_0}{2\mu_0}\,\sqrt{\varepsilon^{;\nu}\varepsilon_{;\nu}}. \tag{13}$$

Up to now, our identification of the rotation tensor $\omega^{\mu\nu}$ of the Elastodynamics of the Spacetime Continuum with the electromagnetic field-strength tensor $F^{\mu\nu}$ has generated consistent results, with no contradictions.

2.3 The Lorentz condition

The Lorentz condition can be derived directly from the theory. Taking the divergence of (4), we obtain

$$A^{\mu}{}_{;\mu} = -\frac{1}{2}\,\varphi_0 u_{\perp}{}^{\mu}{}_{;\mu}. \tag{14}$$

From (23) of [11], viz.

$$\omega^{\mu}{}_{\mu} = u_{\perp}{}^{\mu}{}_{;\mu} = 0, \tag{15}$$

(14) simplifies to

$$A^{\mu}{}_{;\mu} = 0. \tag{16}$$

The Lorentz condition is thus obtained directly from the theory. The reason for the value of zero is that transverse displacements are massless because such displacements arise from a change of shape (distortion) of the spacetime continuum, not a change of volume (dilatation).

2.4 Four-vector potential

Substituting (4) into (5) and rearranging terms, we obtain the equation

$$\nabla^2 A^{\nu} - A^{\mu;\nu}{}_{\mu} = \varphi_0 \omega^{\mu\nu}{}_{;\mu} \tag{17}$$

and, using (3) and (9), this equation becomes

$$\nabla^2 A^{\nu} - A^{\mu;\nu}{}_{\mu} = \mu_{em} j^{\nu}. \tag{18}$$

Interchanging the order of partial differentiation in the term $A^{\mu;\nu}{}_{\mu}$ and using the Lorentz condition of (16), we obtain the well-known wave equation for the four-vector potential [14, see pp. 42–43]

$$\nabla^2 A^{\nu} = \mu_{em} j^{\nu}. \tag{19}$$

The results we obtain are thus consistent with the macroscopic theory of Electromagnetism, with no contradictions.

3 Electromagnetism and the volume force X^{ν}

We now investigate the impact of the volume force X^{ν} on the equations of Electromagnetism. Recalling (7), Maxwell's equation in terms of the rotation tensor is given by

$$\omega^{\mu\nu}{}_{;\mu} = \frac{2\mu_0 + \lambda_0}{2\mu_0}\,\varepsilon^{;\nu} + \frac{1}{2\mu_0}\,X^{\nu}. \tag{20}$$

Substituting for $\omega^{\mu\nu}$ from (3), this equation becomes

$$F^{\mu\nu}{}_{;\mu} = \varphi_0 \frac{2\mu_0 + \lambda_0}{2\mu_0}\,\varepsilon^{;\nu} + \frac{\varphi_0}{2\mu_0}\,X^{\nu}. \tag{21}$$

The additional X^{ν} term can be allocated in one of two ways:

1. either j^ν remains unchanged as given by (10) and the expression for $F^{\mu\nu}{}_{;\mu}$ has an additional term as developed in Section 3.1 below,
2. or $F^{\mu\nu}{}_{;\mu}$ remains unchanged as given by (9) and the expression for j^ν has an additional term as developed in Section 3.2 below.

Option 2 is shown in the following derivation to be the logically consistent approach.

3.1 j^ν unchanged (contradiction)

Using (10) (j^ν unchanged) into (21), Maxwell's equation becomes

$$F^{\mu\nu}{}_{;\mu} = \mu_{em} j^\nu + \frac{\varphi_0}{2\mu_0} X^\nu. \tag{22}$$

Using (20) into (17) and making use of the Lorentz condition, the wave equation for the four-vector potential becomes

$$\nabla^2 A^\nu - \frac{\varphi_0}{2\mu_0} X^\nu = \mu_{em} j^\nu. \tag{23}$$

In this case, the equations for $F^{\mu\nu}{}_{;\mu}$ and A^ν both contain an additional term proportional to X^ν.

We show that this option is not logically consistent as follows. Using (10) into the continuity condition for the current density [14]

$$\partial_\nu j^\nu = 0 \tag{24}$$

yields the expression

$$\nabla^2 \varepsilon = 0. \tag{25}$$

This equation is valid in the macroscopic case where $X^\nu = 0$, but disagrees with the general case (non-zero X^ν) given by (35) of [11], viz.

$$(2\mu_0 + \lambda_0)\nabla^2 \varepsilon = -X^\nu{}_{;\nu}. \tag{26}$$

This analysis leads to a contradiction and consequently is not valid.

3.2 $F^{\mu\nu}{}_{;\mu}$ unchanged (logically consistent)

Proper treatment of the general case requires that the current density four-vector be proportional to the RHS of (21) as follows ($F^{\mu\nu}{}_{;\mu}$ unchanged):

$$\mu_{em} j^\nu = \varphi_0 \frac{2\mu_0 + \lambda_0}{2\mu_0} \varepsilon^{;\nu} + \frac{\varphi_0}{2\mu_0} X^\nu. \tag{27}$$

This yields the following general form of the current density four-vector:

$$j^\nu = \frac{1}{2} \frac{\varphi_0}{\mu_{em}\mu_0} [(2\mu_0 + \lambda_0)\varepsilon^{;\nu} + X^\nu]. \tag{28}$$

Using this expression in the continuity condition for the current density given by (24) yields (26) as required.

Using (28) into (21) yields the same covariant form of the Maxwell equations as in the macroscopic case:

$$F^{\mu\nu}{}_{;\mu} = \mu_{em} j^\nu \tag{29}$$

and the same four-vector potential equation

$$\nabla^2 A^\nu = \mu_{em} j^\nu \tag{30}$$

in the Lorentz gauge.

3.3 Homogeneous Maxwell equation

The validity of this analysis can be further demonstrated from the homogeneous Maxwell equation [14]

$$\partial^\alpha F^{\beta\gamma} + \partial^\beta F^{\gamma\alpha} + \partial^\gamma F^{\alpha\beta} = 0. \tag{31}$$

Taking the divergence of this equation over α,

$$\partial_\alpha \partial^\alpha F^{\beta\gamma} + \partial_\alpha \partial^\beta F^{\gamma\alpha} + \partial_\alpha \partial^\gamma F^{\alpha\beta} = 0. \tag{32}$$

Interchanging the order of differentiation in the last two terms and making use of (29) and the antisymmetry of $F^{\mu\nu}$, we obtain

$$\nabla^2 F^{\beta\gamma} + \mu_{em}(j^{\beta;\gamma} - j^{\gamma;\beta}) = 0. \tag{33}$$

Substituting for j^ν from (28),

$$\nabla^2 F^{\beta\gamma} = -\frac{\varphi_0}{2\mu_0} [(2\mu_0 + \lambda_0)(\varepsilon^{;\beta\gamma} - \varepsilon^{;\gamma\beta}) + (X^{\beta;\gamma} - X^{\gamma;\beta})]. \tag{34}$$

(42) of [11], viz.

$$\mu_0 \nabla^2 \varepsilon^{\mu\nu} + (\mu_0 + \lambda_0)\varepsilon^{;\mu\nu} = -X^{(\mu;\nu)} \tag{35}$$

shows that $\varepsilon^{;\mu\nu}$ is a symmetrical tensor. Consequently the difference term $(\varepsilon^{;\beta\gamma} - \varepsilon^{;\gamma\beta})$ disappears and (34) becomes

$$\nabla^2 F^{\beta\gamma} = -\frac{\varphi_0}{2\mu_0} (X^{\beta;\gamma} - X^{\gamma;\beta}). \tag{36}$$

Expressing $F^{\mu\nu}$ in terms of $\omega^{\mu\nu}$ using (3), the resulting equation is identical to (39) of [11], viz.

$$\mu_0 \nabla^2 \omega^{\mu\nu} = -X^{[\mu;\nu]} \tag{37}$$

confirming the validity of this analysis of Electromagnetism including the volume force.

(28) to (30) are the self-consistent electromagnetic equations derived from the Elastodynamics of the Spacetime Continuum with the volume force. In conclusion, Maxwell's equations remain unchanged. The current density four-vector is the only quantity affected by the volume force, with the addition of a second term proportional to the volume force. It is interesting to note that the current density obtained from the quantum mechanical Klein-Gordon equation with an electromagnetic field also consists of the sum of two terms [16, see p. 35].

4 Discussion and conclusion

In this paper, we have derived Electromagnetism from the Elastodynamics of the Spacetime Continuum based on the identification of the theory's antisymmetric rotation tensor $\omega^{\mu\nu}$ with the electromagnetic field-strength tensor $F^{\mu\nu}$.

The theory provides a physical explanation of the electromagnetic potential: it arises from transverse (shearing) displacements of the spacetime continuum, in contrast to mass which arises from longitudinal (dilatational) displacements of the spacetime continuum. Hence sheared spacetime is manifested as electromagnetic potentials and fields.

In addition, the theory provides a physical explanation of the current density four-vector: it arises from the 4-gradient of the volume dilatation of the spacetime continuum. A corollary of this relation is that massless (transverse) waves cannot carry an electric charge or produce a current.

The transverse mode of propagation involves no volume dilatation and is thus massless. Transverse wave propagation is associated with the distortion of the spacetime continuum. Electromagnetic waves are transverse waves propagating in the *STC* itself, at the speed of light.

The Lorentz condition is obtained directly from the theory. The reason for the value of zero is that transverse displacements are massless because such displacements arise from a change of shape (distortion) of the spacetime continuum, not a change of volume (dilatation).

In addition, we have obtained a generalization of Electromagnetism for the situation where a volume force is present, in the general non-macroscopic case. Maxwell's equations are found to remain unchanged, but the current density has an additional term proportional to the volume force X^{ν}.

The Elastodynamics of the Spacetime Continuum thus provides a unified description of the spacetime deformation processes underlying general relativistic Gravitation and Electromagnetism, in terms of spacetime continuum displacements resulting from the strains generated by the energy-momentum stress tensor.

Submitted on December 9, 2012 / Accepted on December 10, 2012

References

1. Horie K. Geometric Interpretation of Electromagnetism in a Gravitational Theory with Space–Time Torsion. arXiv: hep-th/9409018v1.
2. Sidharth B.G. The Unification of Electromagnetism and Gravitation in the Context of Quantized Fractal Space Time. arXiv: gen-ph/0007021v1.
3. Wu N. Unification of Electromagnetic Interactions and Gravitational Interactions. arXiv: hep-th/0211155v1.
4. Rabounski D. A Theory of Gravity Like Electrodynamics. *Progress in Physics*, 2005, v. 2, 15–29.
5. Wanas M.I. and Ammar S.A. Space-Time Structure and Electromagnetism. arXiv: gr-qc/0505092v2.
6. Shahverdiyev S.S. Unification of Electromagnetism and Gravitation in the Framework of General Geometry. arXiv: gen-ph/0507034v1.
7. Chang Y.-F. Unification of Gravitational and Electromagnetic Fields in Riemannian Geometry. arXiv: gen-ph/0901.0201v1.
8. Borzou A. and Sepangi H.R. Unification of Gravity and Electromagnetism Revisited. arXiv: gr-qc/0904.1363v3.
9. Chernitskii A.A. On Unification of Gravitation and Electromagnetism in the Framework of a General-Relativistic Approach. arXiv: gr-qc/0907.2114v1.
10. Millette P.A. On the Decomposition of the Spacetime Metric Tensor and of Tensor Fields in Strained Spacetime. *Progress in Physics*, 2012, v. 4, 5–8.
11. Millette P.A. The Elastodynamics of the Spacetime Continuum as a Framework for Strained Spacetime. *Progress in Physics*, 2013, v. 1, 55–59.
12. Baylis W.E. Electrodynamics, A Modern Approach. Birkhäuser, Boston, 2002.
13. Jackson J.D. Classical Electrodynamics, 2nd ed. John Wiley & Sons, New York, 1975.
14. Charap J.M. Covariant Electrodynamics, A Concise Guide. The John Hopkins University Press, Baltimore, 2011.
15. Barut A.O. Electrodynamics and Classical Theory of Fields and Particles. Dover Publications, New York, 1980.
16. Greiner W. Relativistic Quantum Mechanics, Wave Equations. Springer-Verlag, New York, 1994.

Peculiar Relations in Cosmology

U. V. S. Seshavatharam* and S. Lakshminarayana†

*Honorary faculty, I-SERVE, Alakapuri, Hyderabad, AP, India. E-mail: seshavatharam.uvs@gmail.com
†Department of Nuclear Physics, Andhra University, Visakhapatnam-03, AP, India. E-mail: lnsrirama@yahoo.com

Within the expanding cosmic Hubble volume, the Hubble length can be considered as the gravitational or electromagnetic interaction range. The product of 'Hubble volume' and 'cosmic critical density' can be called the 'Hubble mass'. Based on this cosmic mass unit, the authors noticed three peculiar semi empirical applications. With these applications it is possible to say that in atomic and nuclear physics, there exists a cosmological physical variable. By observing its rate of change, the future cosmic acceleration can be verified, time to time Hubble's constant can be estimated and finally a unified model of the four cosmological interactions can be developed.

1 Introduction

If we write $R_0 \cong (c/H_0)$ as a characteristic cosmic Hubble radius then the characteristic cosmic Hubble volume is $V_0 \cong \frac{4\pi}{3}R_0^3$. With reference to the critical density $\rho_c \cong \frac{3H_0^2}{8\pi G}$ and the characteristic cosmic Hubble volume, the characteristic cosmic Hubble mass can be expressed as $M_0 \cong \rho_c \cdot V_0 \cong \frac{c^3}{2GH_0}$. If we do not yet know whether the universe is spatially closed or open, then the idea of Hubble volume [1–3] or Hubble mass can be used as a tool in cosmology and unification. This idea is very close to Mach's idea of distance cosmic background. It seems to be a quantitative description to Mach's principle. In understanding the basic concepts of unification of the four cosmological interactions, the cosmic radius (c/H_0) can be considered as the infinite range of the gravitational or electromagnetic interaction. Within the Hubble volume it is noticed that: 1) Each and every point in free space is influenced by the Hubble mass. 2) Hubble mass plays a vital role in understanding the properties of electromagnetic and nuclear interactions and 3) Hubble mass plays a key role in understanding the geometry of the universe.

2 Application 1

Note that large dimensionless constants and compound physical constants reflect an intrinsic property of nature [4,5]. If $\rho_c c^2$ is the present cosmic critical energy density and aT_0^4 is the present cosmic thermal energy density, with this M_0 it is noticed that $\ln\sqrt{\frac{aT_0^4}{\rho_c c^2}\cdot\frac{4\pi\epsilon_0 GM_0^2}{e^2}} \cong \frac{1}{\alpha}$ and at present if $T_0 \cong 2.725$ °K, obtained $H_0 \cong 71.415$ km/sec/Mpc [6,7]. It is also noticed that $\ln\left[\frac{\rho_m}{\rho_c}\sqrt{\frac{4\pi\epsilon_0 GM_0^2}{e^2}}\right] \cong \frac{1}{\alpha}$ where ρ_m is the present cosmic matter density. Obtained $\rho_m \cong 6.70\times10^{-29}$ kg/m^3 is matching with the matter density of spiral and elliptical galaxies. Please note that almost (60 to 70)% of the galaxies are in the form of elliptical and spiral galaxies.

3 Application 2

With this M_0 it is noticed that, $\frac{\hbar c}{Gm_p\sqrt{M_0 m_e}} \cong 1$ where m_p and m_e are the rest masses of proton and electron respectively. This is a very peculiar result. With this relation, obtained value of the present Hubble's constant is 70.75 km/sec/Mpc. From this relation it is clear that, in the presently believed atomic and nuclear "physical constants", there exists one cosmological variable! By observing its cosmological rate of change, the "future" cosmic acceleration can be verified.

4 Application 3

With reference to the Planck mass $M_p \cong \sqrt{\hbar c/G}$ and the elementary charge e, a new mass unit $M_C \cong \sqrt{e^2/4\pi\epsilon_0 G}$ can be constructed. With M_0 and M_C it can be assumed that cosmic thermal energy density, matter energy density and the critical energy density are in geometric series and the geometric ratio is $1+\ln\left(\frac{M_0}{M_C}\right)$. Thus, $\left(\frac{\rho_c c^2}{\rho_m c^2}\right)_0 \cong \left[1+\ln\left(\frac{M_0}{M_C}\right)\right]$ and $\left(\frac{\rho_c c^2}{aT^4}\right)_0 \cong \left[1+\ln\left(\frac{M_0}{M_C}\right)\right]^2$. It is another peculiar observation and the corresponding present CMBR temperature is $T_0 \cong 2.718$ °K. Independent of the cosmic redshift and CMBR observations, with these coincidences it is possible to understand and decide the cosmic geometry. The mystery can be resolved only with further research, analysis, discussions and encouragement.

Submitted on December 13, 2012 / Accepted on December 17, 2012.

References

1. Seshavatharam U. V. S. Physics of rotating and expanding black hole universe. *Progress in Physics*, 2010, v. 2, 7–14.
2. Seshavatharam U. V. S. and Lakshminarayana S. Hubble volume and the fundamental interactions. *International Journal of Applied and Natural Sciences (IJANS)*, 2012, v. 1 (2), 45–58.
3. Seshavatharam U. V. S. and S. Lakshminarayana S. Atom, universe and the fundamental interactions. *Global Journal of Science Frontier Research (A)*, 2012, v. 12 (5), 1.
4. Dirac P. A. M. The cosmological constants. *Nature*, 1937, v. 139, 323.
5. Dirac P. A. M. A new basis for cosmology. *Proc. Roy. Soc. A*, 1938, v. 165, 199.
6. Huchara J. Estimates of the Hubble Constant. *Harvard-Smithsonian Center for Astrophysics*, 2010, http://hubble.plot.dat.
7. Freedman W. L. et al. Final Results from the Hubble Space Telescope Key Project to Measure the Hubble Constant. *The Astrophysical Journal*, 2001, v. 553 (1) 47–72.

Change of Measure between Light Travel Time and Euclidean Distances

Yuri Heymann

3 rue Chandieu, 1202 Geneva, Switzerland. E-mail: y.heymann@yahoo.com

The problem of cosmological distances is approached using a method based on the propagation of light in an expanding Universe. From the change of measure between Light Travel Time and Euclidean Distances, a formula is derived to compute distances as a function of redshift. This formula is identical to Mattig's formula (with $q_0 = 1/2$) which is based on Friedmann's equations of general relativity.

1 Introduction

Euclidean Distances were introduced in [1], and it was suggested that Euclidean Distances need to be used in order to derive the galactic density profile which is the evolution of galactic density over time. The LTD (Light Travel Distance) is the distance traversed by a photon between the time it is emitted and the time it reaches the observer, which may be also referred to as the Light Travel Time. We define the Euclidean distance as the equivalent distance that would be traversed by a photon between the time it is emitted and the time it reaches the observer if there were no expansion of the Universe.

In the present study, a time-varying Hubble coefficient in the Euclidean framework is introduced assuming that the Hubble law observed in the LTD framework is still applicable in the Euclidean framework. The model provides a "kinematic age of the Universe" which is purely mathematical as it is a result of the change of measure between LTDs and Euclidean Distances. A proof is made that a flat Hubble constant in the LTD framework (i.e that does not vary with LTD) is equivalent to a second order forward time-varying Euclidean Hubble coefficient in the Euclidean framework.

2 Foundations of the theory

The observed Hubble constant that is commonly referred to in the literature is a measure of space expansion with respect to LTDs. The Euclidean Hubble coefficient is being defined as the space expansion with respect to Euclidean Distances. This is a change of measure considering that the Euclidean Hubble coefficient varies with time such that the Hubble law is still applicable in the Euclidean framework. This leads to the following equation

$$H_i(t) = \frac{\dot{a}}{a}, \tag{1}$$

where H_i is the instantaneous Euclidean Hubble coefficient, $\dot{a}$ is the Universe expansion velocity and a the scale factor

The main postulate of the present study is that the Euclidean Hubble coefficient needs to be used in order to compute the scale factor in metric distances and not on the basis of LTDs, see (1). If we did not compute the scale factor on the basis of metric distances, the equation would fail to work with cosmological redshifts, which are a homothetic transformation for describing the evolution of light wavelength.

The instantaneous Euclidean Hubble coefficient is defined as the rate of expansion in Euclidean metrics at any given point in time along the trajectory of a light ray reaching the observer.

As space between the photon and the observer expands, this expansion is added to the overall distance the photon has to travel in order to reach the observer; therefore, the Euclidean Distance between the photon and the observer is defined by the following differential equations, respectively in the temporal and metric form:

1) In the LTD framework (the temporal form)

$$\frac{dy}{dt} = -c + H_0\, c\, T, \tag{2}$$

where: y is the Euclidean Distance between the photon and the observer, T the LTD between the observer and the photon, c the celerity of light, and H_0 the Hubble constant as of today;

2) In the Euclidean framework (the metric form)

$$\frac{dy}{dt} = -c + H_i(t)\, y, \tag{3}$$

where y is the Euclidean Distance between the photon and the observer, c the celerity of light, and $H_i(t)$ the Euclidean time-varying Hubble coefficient.

For the purpose of convenience let us consider the following form for the Euclidean time-varying Hubble coefficient

$$H_i(t) = \frac{n}{t}, \tag{4}$$

where $H_i(t)$ is the Euclidean time-varying Hubble coefficient, n the order of the time-varying Euclidean Hubble, and t the time from the hypothetical big bang for which time was set to zero.

Note that in the present study both the Hubble constant and the Euclidean Hubble coefficient are expressed in units of $[time^{-1}]$ by converting all distances into Light Travel Time, and with the celerity $c = 1$.

3 Proof that a flat Hubble constant in the LTD framework is time varying of order two in the Euclidean framework

First, let us solve the differential equation for the propagation of light in the LTD framework assuming a flat Hubble constant (i.e. that does not vary with LTD). Let us consider a photon initially situated at a Euclidean Distance y_0 from the observer and moving at celerity c in the direction of the observer. Let us say T is the initial LTD between the photon and the observer, and define the Hubble constant function of LTDs.

The differential equation describing the propagation of light in the LTD framework is described by (2). By setting time zero at a reference T_b in the past, we have $t = T_b - T$; therefore, $dt = -dT$. Hence, (2) becomes

$$\frac{dy}{dT} = c - H_0\, c\, T\,, \tag{5}$$

with boundary conditions $y(T) = y_0$, and $y(0) = 0$.

By integration from 0 to T, the following relationship relating Euclidean Distances y to Light Travel Distances T is obtained

$$y = c\, T - \frac{c\, H_0\, T^2}{2}\,. \tag{6}$$

Now let us derive the differential equation for the propagation of light in the Euclidean framework assuming the time-varying Hubble coefficient from (4) (see Figure 1). From the differential equation describing the propagation of light in the Euclidean framework (3), we get

$$\frac{dy}{dt} = -c + \frac{n}{t}\, y\,. \tag{7}$$

By integrating this first order non-homogeneous differential equation between $T_b - T$ and T_b, the following solution is obtained which describes the relationship between Euclidean Distances and LTDs

$$y = \frac{c}{n-1}\left(Tb - T - Tb^{1-n}\,(Tb - T)^n\right). \tag{8}$$

By setting n equal to 2 in (8) for a second order time-varying Hubble coefficient, we get

$$y = c\left(T - \frac{T^2}{T_b}\right). \tag{9}$$

Based on the recession speed, the relationship between the Hubble constant defined function of LTDs, and the Euclidean Hubble, for T small is as follows

$$H_0\, c\, T = \frac{n}{(Tb - T)}\, y\,. \tag{10}$$

Hence, $\frac{n}{T_b}$ is obtained by computing the following limit

$$\frac{n}{T_b} = \lim_{T\to 0}\left(\frac{H_0\, c\, T}{y}\right). \tag{11}$$

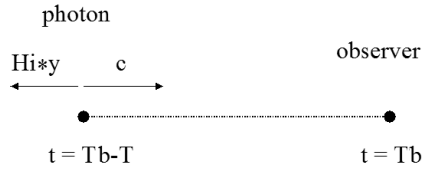

Fig. 1: Schema to represent the propagation of light in an expanding space in the Euclidean framework. Where T is the Light Travel Distance between the observer and the source of light, T_b is the kinematic age of the Universe, and n the order of the time-varying Hubble coefficient (time zero set at time T_b from today).

By substitution of y from (8), we get

$$\frac{n}{T_b} = \lim_{T\to 0}\left(\frac{(n-1)\, T\cdot H_0}{T_b - T - T_b^{1-n}\,(T_b - T)^n}\right) = H_0\,. \tag{12}$$

Therefore, the "kinematic age of the Universe" is

$$T_b = \frac{n}{H_0}, \tag{13}$$

with H_0 the Hubble constant as of today.

By substitution of $T_b = \frac{2}{H_0}$ into (9), we get

$$y = c\, T - \frac{c\, H_0\, T^2}{2}\,. \tag{14}$$

This solution is identical to (6) relating LTDs to Euclidean Distances for the flat Hubble constant in the LTD framework. This is the proof that a flat Hubble constant in the LTD framework is equivalent to a time-varying Hubble coefficient of order two in the Euclidean framework. The equation $H_i(t) = 2/t$ is the connection between (2) and (3).

We can easily show that the recession speed with the second order time-varying Hubble coefficient in the Euclidean framework is the same as the recession speed in the LTD framework. The calculations are as follow

$$H_i(t)\; y = \frac{2}{t}\, y = \frac{2\, c}{T_b - T}\left(T - \frac{H_0\, T^2}{2}\right). \tag{15}$$

By substitution of T_b from (13) (with a second order time-varying Hubble coefficient) into (15), we obtain

$$H_i(t)\, y = H_0\, c\, T\,, \tag{16}$$

where T is the LTD between the observer and the source of light, and y the Euclidean Distance.

4 Evolutionary model of the scale factor

The differential equation describing the evolution of the scale factor a is as follows, identical to (1),

$$\frac{da}{dt} = H_i(t)\, a\,. \tag{17}$$

As $H_i(t) = \frac{2}{t}$, we get

$$\int_{a_1}^{a_0} \frac{1}{a} da = \int_{T_b - T}^{T_b} \frac{2}{t} dt . \quad (18)$$

By integrating (18) , we obtain

$$\ln\left(\frac{a_0}{a_1}\right) = 2 \ln\left(\frac{T_b}{T_b - T}\right), \quad (19)$$

which is equivalent to

$$\frac{a_0}{a_1} = \left(\frac{T_b}{T_b - T}\right)^2 . \quad (20)$$

5 Expression of distances versus redshifts

From cosmological redshifts, we have

$$1 + z = \frac{a_0}{a_1} , \quad (21)$$

where a_0 is the present scale factor, a_1 the scale factor at redshift z.

Combining (20) and (21), we get

$$T = T_b \left(1 - \frac{1}{\sqrt{1+z}}\right). \quad (22)$$

By substitution of T_b from (13) for a second order time-varying Hubble coefficient, we get the following equation relating LTD to redshifts

$$T = \frac{2}{H_0} \left(1 - \frac{1}{\sqrt{1+z}}\right). \quad (23)$$

6 Comparison with the equation of Mattig

The equation of Mattig [2] is as follows

$$rR_0 = \frac{1}{H_0 q_0^2 (1+z)} \times \times \left(q_0 z + (q_0 - 1)(\sqrt{1 + 2q_0 z} - 1)\right), \quad (24)$$

where r is the distance, q_0 is the deceleration parameter, R_0 the present scale factor, z the redshift, H_0 the present scale factor.

For comparison purpose with the equation of the present study, we should set q_0 equal to 1/2 (flat matter dominated Universe), and R_0 to 1. Therefore, we obtain

$$r = \frac{2}{H_0} \left(1 - \frac{1}{\sqrt{1+z}}\right). \quad (25)$$

This formula is identical to (23). We have just shown that the solution to our problem is identical to Mattig formula for q_0 equal to 1/2.

7 Discussion

Based on the change of measure between LTD and Euclidean Distances, a formula that expresses distances versus redshifts is obtained. From the change of framework between LTD and Euclidean distances, it has been proved that a flat Hubble constant (that does not vary with LTD) is equivalent to a time-varying Euclidean Hubble coefficient of order two. Finally, the evolutionary model of the scale factor is derived and matched to the cosmological redshift equation in order to obtain the LTD versus redshift equation. This equation is identical to Mattig's formula (with $q_0 = 1/2$) which is based on Friedmann's equations of general relativity. The Euclidean Hubble coefficient was used in order to derive the evolution of the scale factor in metric distances; otherwise, the cosmological redshift equation would not be applicable to light wavelengths. This study proposes a new approach to compute cosmological distances which is based on the introduction of Euclidean Distances in addition to Light Travel Distances in an expanding Universe, and a change of measure. The calculations involved are quite simple and our definition of Euclidean Distances may be used as a source of inspiration to develop future cosmological models.

Acknowledgements

The author is deeply indebted to Edmund Wood for his support in correcting my English.

Submitted on December 4, 2012 / Accepted on December 15, 2012

References

1. Heymann Y. Building galactic density profiles. *Progress in Physics*, 2011, v. 4, 63–67.
2. Mattig W. Über den Zusammenhang zwischen Rotverschiebung und scheinbarer Helligkeit. *Astronomische Nachrichten*, 1958, v. 284, 109–111.

E(5) Characters to ^{100}Ru Isotope

Sohair M. Diab* and Salah A. Eid†
*College of Sciences (Women Section), Phys. Dept., Umm Al-Qura University, Saudi Arabia.
†Faculty of Engineering, Phys. Dept., Ain Shams University, Cairo, Egypt. E-mail: mppe2@yahoo.co.uk

The positive and negative parity states of ^{100}Ru isotope are studied within the frame work of the interacting boson approximation model ($IBA-1$). The calculated levels energy, potential energy surfaces, $V(\beta,\gamma)$, and the electromagnetic transition probabilities, $B(E1)$ and $B(E2)$, show that ^{100}Ru isotope has $E(5)$ Characters. Staggering effect, $\Delta I = 1$, has been observed between the positive and negative parity states. The electric monopole strength, $X(E0/E2)$, has been calculated. All calculated values are compared to the available experimental, theoretical data and reasonable agreement has been obtained.

1 Introduction

The mass region $A = 100$ has been of considerable interst for nuclear structure studies as it shows many interesting features. These nuclei show back bending at high spin and shape transitions from vibrational to γ-soft and rotational characters. Many attempts have made to explore these structural changes which is due mainly to the n-p interactions.

Experimentally, the nuclear reaction ^{100}Mo (α,xn) [1] has been used in studying levels energy of ^{100}Ru. Angular distribution, γ-γ coincidences were measured, half-life time has calculated and changes to the level scheme were proposed. Also, double beta decay rate of ^{100}Mo to the first excited 0^+ state of ^{100}Ru has been measured experimentally [2] using γ-γ coincidence technique.

Doppler-shift attenuation measurements following the ^{100}Ru (n, $n'\gamma$) reaction [3] has used to measure the life times of the excited states in ^{100}Ru. Absolute transition rates were extracted and compared with the interacting boson model description. The 2^+(2240.8 keV) state which decays dominantly to the 2^+ via 1701 keV transition which is almost pure $M1$ in nature considered as a mixed-symmetry state. Again ^{100}Ru has been studied [4] experimently and several levels were seen where some new ones are detected below 3.2 MeV.

Theoretically many models have been applied to ruthenium isotopes. Yukawa folded mean field [5] has applied to ^{100}Ru nucleus while the microscopic vibrational model has applied to ^{104}Ru and some other nuclei with their daughters [6]. The latter model was successful in describing the yrast 0^+ and 2^+ states of most of these nuclei and also some of their half-lives.

The very high-spin states of nuclei near A$\approx$100 are investigated by the Cranked Strutinsky method [7] and many very extended shape minima are found in this region. Interacting boson model has been used in studying *Ru* isotopes using a $U(5)$–$O(6)$ transitional Hamiltonian with fixed parameters [8-10] except for the boson number N. Hartree-Fock Bogoliubov [11] wave functions have been tested by comparing the theoretically calculated results for ^{100}Mo and ^{100}Ru nuclei with the available experimental data. The yrast spectra, reduced $B(E2, 0^+ \rightarrow 2^+)$ transition probabilities, quadrupole moments $Q(2^+)$ and g factors, g(2^+) are computed. A reasonable agreement between the calculated and observed values has been obtained.

The microscopic anharmonic vibrator approach (MAVA) [12] has been used in investigating the low-lying collective states in $^{98-108}$Ru. Analysis for the level energies and electric quadrupole decays of the two-phonon type of states indicated that ^{100}Ru can interpreted as being a transitional nucleus between the spherical anharmonic vibrator ^{98}Ru and the quasirotational $^{102-106}$Ru isotopes.

A new emprical approach has been proposed [13] which is based on the connection between transition energies and spin. It allows one to distinguish vibrational from rotational characters in atomic nuclei. The cranked interacting boson model [14] has been used in estimating critical frequencies for the rotation-induced spherical-to-deformed shape transition in $A = 100$ nuclei. The predictions show an agreement with the back bending frequencies deduced from experimental yrast sequences in these nuclei.

The aim of the present work is to use the $IBA-1$ [15, 16] for the following tasks:

1. Calculating the potential energy surfaces, $V(\beta,\gamma)$, to know the type of deformation existing for ^{100}Ru;
2. Calculating levels energy, electromagnetic transition rates $B(E1)$ and $B(E2)$;
3. Studying the relation between the angular momentum I and the rotational angular frequency $\hbar\omega$ for bending in ^{100}Ru;
4. Calculating staggering effect to see beat patterns and detect any interactions between the (+ve) and (−ve) parity states, and
5. Calculating the electric monopole strength $X(E0/E2)$.

2 (IBA-1) model

2.1 Level energies

IBA-1 model was applied to the positive and negative parity states of ^{100}Ru isotope. The Hamiltonian employed in the

nucleus	EPS	$PAIR$	ELL	QQ	OCT	HEX	$E2SD(eb)$	$E2DD(eb)$
^{100}Ru	0.5950	0.000	0.0085	−0.0200	0.0000	0.0000	0.1160	−0.3431

Table 1: Parameters used in IBA-1 Hamiltonian (all in MeV).

present calculation is:

$$\begin{aligned} H = \; & EPS \cdot n_d + PAIR \cdot (P \cdot P) \\ & + \tfrac{1}{2}\, ELL \cdot (L \cdot L) + \tfrac{1}{2}\, QQ \cdot (Q \cdot Q) \\ & + 5\, OCT \cdot (T_3 \cdot T_3) + 5 HEX \cdot (T_4 \cdot T_4) \end{aligned} \tag{1}$$

where

$$P \cdot p = \frac{1}{2} \left[\begin{array}{l} \left\{ (s^\dagger s^\dagger)^{(0)}_0 - \sqrt{5} (d^\dagger d^\dagger)^{(0)}_0 \right\} x \\ \left\{ (ss)^{(0)}_0 - \sqrt{5} (\tilde{d}\tilde{d})^{(0)}_0 \right\} \end{array} \right]^{(0)}_0 , \tag{2}$$

$$L \cdot L = -10 \sqrt{3} \left[(d^\dagger \tilde{d})^{(1)} x (d^\dagger \tilde{d})^{(1)} \right]^{(0)}_0 , \tag{3}$$

$$Q \cdot Q = \sqrt{5} \left[\begin{array}{l} \left\{ (S^\dagger \tilde{d} + d^\dagger s)^{(2)} - \dfrac{\sqrt{7}}{2} (d^\dagger \tilde{d})^{(2)} \right\} x \\ \left\{ (s^\dagger \tilde{d} + + \tilde{d} s)^{(2)} - \dfrac{\sqrt{7}}{2} (d^\dagger \tilde{d})^{(2)} \right\} \end{array} \right]^{(0)}_0 , \tag{4}$$

$$T_3 \cdot T_3 = - \sqrt{7} \left[(d^\dagger \tilde{d})^{(2)} x (d^\dagger \tilde{d})^{(2)} \right]^{(0)}_0 , \tag{5}$$

$$T_4 \cdot T_4 = 3 \left[(d^\dagger \tilde{d})^{(4)} x (d^\dagger \tilde{d})^{(4)} \right]^{(0)}_0 . \tag{6}$$

In the previous formulas, n_d is the number of boson; $P \cdot P$, $L \cdot L$, $Q \cdot Q$, $T_3 \cdot T_3$ and $T_4 \cdot T_4$ represent pairing, angular momentum, quadrupole, octupole and hexadecupole interactions between the bosons; EPS is the boson energy; and $PAIR$, ELL, QQ, OCT, HEX is the strengths of the pairing, angular momentum, quadrupole, octupole and hexadecupole interactions.

2.2 Transition rates

The electric quadrupole transition operator employed in this study is:

$$\begin{aligned} T^{(E2)} = \; & E2SD \cdot (s^\dagger \tilde{d} + d^\dagger s)^{(2)} + \\ & + \tfrac{1}{\sqrt{5}}\, E2DD \cdot (d^\dagger \tilde{d})^{(2)} . \end{aligned} \tag{7}$$

$E2SD$ and $E2DD$ are adjustable parameters.

The reduced electric quadrupole transition rates between $I_i \to I_f$ states are given by:

$$B(E_2, I_i - I_f) = \frac{[< I_f \parallel T^{(E_2)} \parallel I_i >]^2}{2 I_i + 1} . \tag{8}$$

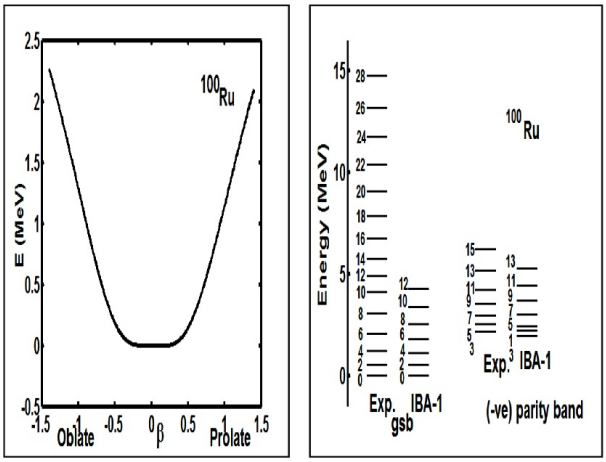

Fig. 1: A: Potential energy surfaces for ^{100}Ru. B: Comparison between exp. [19] and theoretical IBA-1 energy levels.

3 Results and discussion

3.1 The potential energy surfaces

The potential energy surfaces [17], $V(\beta, \gamma)$, as a function of the deformation parameters β and γ are calculated using:

$$\begin{aligned} & E_{N_\Pi N_\nu}(\beta, \gamma) = \; < N_\pi N_\nu; \beta\gamma \,|H_{\pi\nu}| N_\pi N_\nu; \beta\gamma > \; = \\ & = \zeta_d (N_\nu N_\pi) \beta^2 (1 + \beta^2) + \beta^2 (1 + \beta^2)^{-2} \times \\ & \times \left\{ k N_\nu N_\pi [4 - (\bar{X}_\pi \bar{X}_\nu) \beta \cos 3\gamma] \right\} + \\ & + \left\{ [\bar{X}_\pi \bar{X}_\nu \beta^2] + N_\nu (N_\nu - 1) \left(\frac{1}{10}\, c_0 + \frac{1}{7}\, c_2 \right) \beta^2 \right\} , \end{aligned} \tag{9}$$

where

$$\bar{X}_\rho = \left(\frac{2}{7} \right)^{0.5} X_\rho \qquad \rho = \pi \;\; \text{or} \;\; \upsilon . \tag{10}$$

The calculated potential energy surfaces, $V(\beta, \gamma)$, are presented in Fig. 1A. The flat potential in the critical symmetry point has supported quite well the $E(5)$ characters to ^{100}Ru nucleus. Also, the energy ratios presented in Table 4 as well as the existance of ^{100}Ru isotope between the spherical anharmonic vibrator ^{98}Ru and γ - soft ^{102}Ru nuclei [9] supported the $E(5)$ characters.

3.2 Energy spectra

The energy of the positive and negative parity states of ^{100}Ru isotope are calculated using computer code PHINT [18]. A

$I_i^+ I_f^+$	B(E2)	$I_i^+ I_f^+$	B(E1)
0_1 Exp*. 2_1	0.490(5)	$1_1\ 0_1$	0.0030
0_1 Theor. 2_1	0.4853	$1_1\ 0_2$	0.1280
$2_1\ 0_1$	0.0970	$3_1\ 2_1$	0.1211
$2_2\ 0_1$	0.0006	$3_1\ 2_2$	0.0415
$2_2\ 0_2$	0.0405	$3_1\ 2_3$	0.0002
$2_3\ 0_1$	0.0000	$3_2\ 2_1$	0.0024
$2_3\ 0_2$	0.0759	$3_2\ 2_2$	0.0197
$2_3\ 0_3$	0.0087	$3_2\ 2_3$	0.2126
$2_4\ 0_3$	0.0066	$5_1\ 4_1$	0.2533
$2_4\ 0_4$	0.0588	$5_1\ 4_2$	0.0480
$4_1\ 2_1$	0.1683	$5_1\ 4_3$	0.0006
$4_1\ 2_2$	0.0142	$7_1\ 6_1$	0.3950
$4_1\ 2_3$	0.0319	$7_1\ 6_2$	0.0446
$6_1\ 4_1$	0.2039	$9_1\ 8_1$	0.5439
$6_1\ 4_2$	0.0179	$9_1\ 8_2$	0.0342
$6_1\ 4_3$	0.0242	$11_1\ 10_1$	0.6983
$8_1\ 6_1$	0.2032		
$8_1\ 6_2$	0.0183		
$8_1\ 6_3$	0.0157		
$10_1\ 8_1$	0.1678		
$10_1\ 8_2$	0.0175		

Table 2: Calculated $B(E2)$ and $B(E1)$ in ^{100}Ru.

I_i^+	I_f^+	I_f^+	$X_{if'f}\,(E0/E2)^{100}$Ru
0_2	0_1	2_1	0.027
0_3	0_1	2_1	0.347
0_3	0_1	2_2	0.009
0_3	0_1	2_3	0.042
0_3	0_2	2_1	0.086
0_3	0_2	2_2	0.002
0_3	0_2	2_3	0.010
0_4	0_1	2_2	0.010
0_4	0_1	2_3	0.010
0_4	0_1	2_4	0.113
0_4	0_2	2_2	0.030
0_4	0_2	2_3	0.034
0_4	0_2	2_4	0.340
0_4	0_3	2_1	0.454
0_4	0_3	2_2	0.010
0_4	0_3	2_3	0.011
0_4	0_3	2_4	0.113

Table 3: Calculated $X_{if'f}\,(E0/E2)$.

Energy Ratios	$E_{4_1^+}/E_{2_1^+}$	$E_{2_2^+}/E_{2_1^+}$
$E(5)$	2.19	2.20
Exp. [19]	2.27	2.52
$IBA-1$	2.12	2.11

Table 4: Energy ratios of $E(5)$ characters to ^{100}Ru.

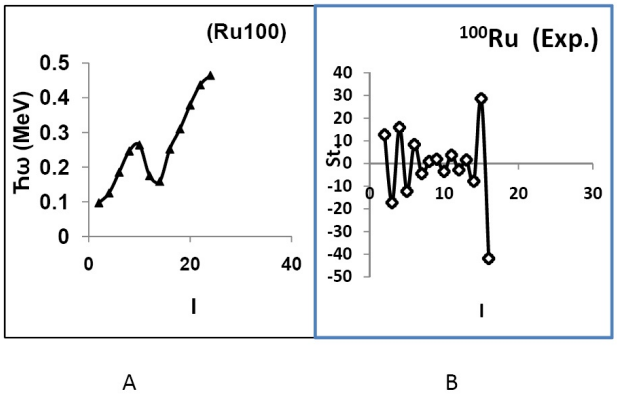

Fig. 2: A: Angular momentum I as a function of $(\hbar\omega)$. B: $(\Delta I = 1)$, staggering pattern for ^{100}Ru isotope.

comparison between the experimental spectra [19] and our calculations, using values of the model parameters given in Table 1 for the ground state band are illustrated in Fig. 1B. The agreement between the calculated levels energy and their correspondence experimental values are slightly higher especially for the higher excited states. We believe this is due to the change of the projection of the angular momentum which is due mainly to band crossing.

Unfortunately there is no enough measurements of electromagnetic transition rates $B(E1)$ or $B(E2)$ for ^{100}Ru nucleus. The only measured $B(E2, 0_1^+ \to 2_1^+)$ is presented, in Table 2 for comparison with the calculated values[20]. The parameters $E2SD$ and $E2DD$ displayed in Table 1 are used in the computer code NPBEM [18] for calculating the electromagnetic transition rates after normalized to the available experimental values. No new parameters are introduced for calculating electromagnetic transition rates $B(E1)$ and $B(E2)$ of intraband and interband.

The moment of inertia I and angular frequency $\hbar\omega$ are calculated using equations (11, 12):

$$\frac{2I}{\hbar^2} = \frac{4I-2}{\Delta E(I \to I-2)}\,, \tag{11}$$

$$(\hbar\omega)^2 = (I^2 - I + 1)\left[\frac{\Delta E(I \to I-2)}{(2I-1)}\right]^2\,. \tag{12}$$

The plot in Fig. 2A show back bending at angular momentum $I^+ = 10$. It means, there is a crossing between the $(+ve)$ and $(-ve)$ parity states in the ground state band which

was confirmed by calculating the staggering effect where a beat pattern has been observed, Fig. 2B.

3.3 Electric monopole transitions

The electric monopole transitions, $E0$, are normally occurring between two states of the same spin and parity by transferring energy and zero unit of angular momentum. The strength of the electric monopole transition, $X_{if'f}(E0/E2)$, [21] can be calculated using equations (13, 14); results are presented in Table 3

$$X_{if'f}(E0/E2) = \frac{B(E0, I_i - I_f)}{B(E2, I_i - I_{'f})}\,, \tag{13}$$

$$X_{if'f}(E0/E2) = (2.54\times 10^9)\,A^{3/4}\times \frac{E_\gamma^5(\mathrm{MeV})}{\Omega_{KL}}\,\alpha(E2)\,\frac{T_e(E0, I_i - I_f)}{T_e(E_2, I_i - I_{'f})}\,. \tag{14}$$

Here: A is mass number; I_i is spin of the initial state where E0 and E2 transitions are depopulating it; I_f is spin of the final state of E0 transition; $I_{'f}$ is spin of the final state of E2 transition; E_γ is gamma ray energy; Ω_{KL} is electronic factor for K,L shells [22]; $\alpha(E2)$ is conversion coefficient of the E2 transition; $T_e(E0, I_i - I_f)$ is absolute transition probability of the $E0$ transition between I_i and I_f states, and $T_e(E_2, I_i - I_{'f})$ is absolute transition probability of the $E2$ transition between I_i and $I_{'f}$ states.

3.4 The staggering

The presence of $(+ve)$ and $(-ve)$ parity states has encouraged us to study staggering effect [23–25] for ^{100}Ru isotope using staggering function equations (15, 16) with the help of the available experimental data [19].

$$St(I) = 6\Delta E(I) - 4\Delta E(I-1) - 4\Delta E(I+1) + \Delta E(I+2) + \Delta E(I-2)\,, \tag{15}$$

with

$$\Delta E(I) = E(I+1) - E(I)\,. \tag{16}$$

The calculated staggering patterns are illustrated in Fig. 2B and show an interaction between the $(+ve)$ and $(-ve)$ parity states for the ground state band of ^{100}Ru.

3.5 Conclusions

IBA-1 model has been applied successfully to ^{100}Ru isotope and:

1. The levels energy are successfully reproduced;
2. The potential energy surfaces are calculated and show $E(5)$ Characters to ^{100}Ru;
3. Electromagnetic transition rates $B(E1)$ and $B(E2)$ are calculated;
4. Bending for ^{100}Ru has been observed at angular momentum $I^+ = 10$;
5. Strength of the electric monopole transitions $X_{if'f}(E0/E2)$ are calculated; and
6. Staggering effect has been calculated and beat pattern has been obtained, showing an interaction between the $(-ve)$ and $(+ve)$ parity states.

Submitted on: December 12, 2012 / Accepted on: December 20, 2012

References

1. Maazek J., Honusek M., Spalek A., Bielcik J., Slivova J. and Pasternak A. A. Levels of ^{100,101}Ru excited in the reaction ^{100}Mo($\alpha\chi$n). *Acta Physica Polonica B*, 1998, v. 29, 433.
2. Braeckeleer L. L., Hornish M., Barabash A. and Umatov V. Measurement of the $\beta\beta$-decay rate of ^{100}Mo to the first excited 0^+ state of ^{100}Ru. *Physical Review*, 2001, v. 86, 3510.
3. Genilloud L., Brown T. B., Corminboeuf F., Garrett P. E., Hannant C. D., Jolie J., Warr N. and Yates S. W. Characterization of the three-phonon region of ^{100}Ru. *Nuclear Physics A*, 2001, v. 683, 287.
4. Horodynski-Matsushigue L. H., Rodrigues C. L., Sampaio F. C. and Lewin T. B. Particle spectroscopy of low-lying collective states in ^{100}Ru. *Nuclear Physics A*, 2002, v. 709, 73.
5. Nerlo-Pomorska B., Pomorski K. and Bartel J. Shell energy and the level- density parameter of hot nuclei. *Physical Review C*, 2006, v. 74, 034327.
6. Raina P. K. and Dhiman R. K. Systematics of $\beta\beta$ decay sensitive medium mass nuclei using quadrupole-quadrupole plus pairing interactions. *Physical Review C*, 2001, v. 64, 024310.
7. Chasman R. R. Very extended nuclear shapes near $A = 100$. *Physical Review C*, 2001, v. 64, 024311.
8. Frank A., Alonso C. E. and Arias J. M. Search for $E(5)$ symmetry in nuclei: the *Ru* isotopes. *Physical Review C*, 2001, v. 65, 014301.
9. Diab S. M. , Eid S. A. Nature of the excited states of the even-even $^{98\text{-}108}$Ru istopes. *Progress in Physics*, 2008, v. 4, 51.
10. Cejnar P., Jolie J. and Kern J. Universal anharmonicity anomly in nuclei. *Physical Review C*, 2001, v. 63, 047304.
11. Chaturvedi K., Dixit B. M. and Rath P. K. Two neutrino double β decay of ^{100}Mo to the 2^+ excited state of ^{100}Ru. *Physical Review C*, 2003, v. 67, 064317.
12. Kotila J., Suhonen J. and Delion D. S. Low-lying collective states in ^{98,106}Ru isotopes studied using a microscopic anharmonic vibrator approach. *Physical Review C*, 2003, v. 68, 054322.
13. Regan P. H., Beausang C. W., Zamfir N. V., Casten R. F., Zhang J., Meyer D. A. and Ressler J. J. Signature for vibrational to rotational evolution along the yrast line. *Physical Review Letters*, 2003, v. 90 , 152502.
14. Cejnar P. and Jolie J. Rotation-driven spherical-to-deformed shape transition in $A \approx 100$ nuclei and the cranked interacting boson model. *Physical Review C*, 2004, v. 69, 011301.
15. Arima A. and Iachello F. Interacting boson model of collective states: The vibrational limit. *Annals of Physics (N.Y.)*, 1976, v. 99, 253.
16. Arima A. and Iachello F. Interacting boson model of collective states: The rotational limit. *Annals of Physics (N.Y.)*, 1978, v. 111, 201.
17. Ginocchio J. N. and Kirson M. W. An intrinisic state for the interacting boson model and its relationship to the Bohr- Mottelson approximation. *Nuclear Physics A*, 1980, v. 350, 31.
18. Scholten O. *The programm package PHINT (1980) version, internal report KVI-63, Gronigen: Keryfysisch Versneller Institut.*
19. Balraj Singh. Adopted levels, gammas for ^{100}Ru. *Nuclear Data Sheets*, 1997, v. 81, 1997.

20. Raman S., Nestor JR.C.W., and Tikkanen P., Transition probability from the ground to the first-excited 2^+ state of even-even nuclides. *Atomic Data and Nuclear Data Tables*, 2001, v. 78, 1.

21. Rasmussen J. O. Theory of $E0$ transitions of spheroidal nuclei. *Nuclear Physics*, 1960, v. 19, 85.

22. Bell A. D ., Aveldo C. E., Davidson M. G. and Davidson J. P., Table of $E0$ conversion probability electronic factors. *Canadian Journal of Physics*, 1970, v. 48, 2542.

23. Minkov N., Yotov P., Drenska S. and Scheid W. Parity shift and beat staggering structure of octupolr bands in a collective model for quadrupole-octupole0deformed nuclei. *Journal of Physics G*, 2006, v. 32, 497.

24. Bonatsos D., Daskaloyannis C., Drenska S. B., Karoussos N., Minkov N., Raychev P. P. and Roussev R. P. $\Delta I = 1$ staggering in octupole bands of light actinides Beat patterns. *Physical Review C*, 2000, v. 62, 024301.

25. Minkov N., Drenska S. B., Raychev P. P., Roussev R. P. and Bonatsos D. Beat patterns for the odd-even staggering in octupole bands from quadrupole-octupole Hamiltonian. *Physical Review C*, 2001, v. 63, 044305.

A New Large Number Numerical Coincidences

Alexander Kritov
E-mail: alex@kritov.ru

In this article, the author gives a set of new hypothesis wherein he presents new, exact and simple relations between physical constants and numbers. The author briefly analyses the discovered coincidences in terms of their accuracy and confidence, while leaving himself aside any physical explanation of the presented formulas. Important: all the found relations have a common nature of the "power of two". The exact nature of this remains unknown for yet, so it requires further research. The presented material may also be viewed as a logical continuation and development of Dirac's and Eddington's Large Numbers Hypothesis (LNH). However, in contrast to Dirac's LNH, two of the presented ratios are not approximate but manifest exact equality. This allows a theoretical prediction of the Universe's radius as well as a calculation of the exact value of Newtonian gravitational constant G, which all fall within the range of the current measurement data and precision. The author formulates these Large Number Numerical (LNN) coincidences by realizing that further discovery of their meaning may lead to a significant change in our understanding of Nature. In this work, SI units are used.

Introduction

Many attempts of bringing together physics and numerology had been done before but a very important step was done in 1938 by Arthur Eddington. According to Eddington's proposal the number of protons in the entire Universe should be exactly equal to: $N_{Edd} = 136 \times 2^{256} \approx 10^{79}$ [1, 2, 17]. So, it was hypothesised that square root of N_{Edd} should be close to Dirac's Big number $N \approx \sqrt{136 \times 2^{256}} = \sqrt{136} \times 2^{128}$. Later on, Eddington changed 136 to 137 and insisted that the fine structure constant has to be precisely 1/137, and then his theory seemed to fail at this cornerstone. However, Eddington's statement also had the number $(2^{128})^2$ which has been left without proper attention. Actually, few years earlier, in 1929, it was German physicist R. Fürth who proposed to use 16^{32} (which is also 2^{128}) in order to connect gravitation to atomic constants [10]. However, all these coincidences have been left unexplained until present time. As G. Gamov said [16]: "Since the works of Sir Arthur Eddington, it has become customary to discuss from time to time the numerical relations between various fundamental constants of nature". For example, another interesting attempt to use "a log-base-2 relation" between electromagnetic and gravitational coupling constant was made by Saul-Paul Sirag, the researcher from San Francisco in 1979 [12]. Particularly, as noted, power of 2 should have significant role in numerical relations for physics constants according to the author's idea.

Suggested four Large Number Numerical (LNN) relations or coincidences are presented below. These coincidences are not dependent and related to each other, so prove or disprove of one of them does not mean the same for the others. They all have common number of 2^{128}. First two relations seem to be exact equations, and second two are valid with defined uncertainty. Because of that their nature is more hypothetical, so second two relations are also called "weak".

1 Cosmological coincidence

The relation is analogous to famous Dirac's ratio $R_U/r_e \sim 10^{40}$ which relates the Universe radius with classical electron radius. However, Dirac's ratio is actually valid only approximately (with precision of "the same order of magnitude"), in opposite, the suggested replacement is an exact equation given as follows:

$$\frac{R_U}{\lambda_e} = 2^{128}, \tag{1}$$

where R_U is value for the radius of the observable Universe and $\lambda_e = \hbar/m_e c \approx 3.86 \times 10^{-13}$ (m) is electron's reduced Compton wavelength (De Broglie wave). The relation (1) provides us with precise size and age of the observable Universe. So it leads to exact value for the Universe radius of $R_U = 1.314031 \times 10^{26}$ meters corresponding to the Universe age of 13.8896 billion years.

Recently F.M.Sanchez, V.Kotov, C.Bizouard discovered that the use of the reduced electron Compton wavelength is decisive for the compatibility of the Hubble-Lemaitre length with 2^{128} [13–15]. They use this length unit because of proposed holographic relation involving it. Here, the author independently developes this idea suggesting that (1) is an exact relation.

The age of the Universe, according to the Wilkinson Microwave Anisotropy Probe (WMAP) 7-year results, is 13.75 ± 0.13 billion years [9]. Latest NASA observation by Hubble gives the age of the Universe as 13.7 billion years [3]. It is very close to the obtained value and lies in the existing error range. So, the coincidence (1) seems to define the exact Universe elapsed life time as:

$$T_U = \frac{\lambda_e}{c} 2^{128}. \tag{1.1}$$

Important to note, that having (1.1), initial Dirac's relation may be expressed in the following form:

$$N_1 = \frac{R_U}{r_e} = \alpha^{-1}\, 2^{128}, \tag{1.2}$$

where $\alpha^{-1} = 137.036$ is inverted fine structure constant and $r_e = ke^2/(m_e c^2)$ — classical electron radius with eliminated numerical factor (i.e. equal to unity) and N_1 is exact value for the large number introduced by Dirac (4.66×10^{40}). As we know for sure that the Universe is expanding and $R_U(t)$ is dependent on time, so the equation (1) suggest that one or few of the fundamental constants (h, c, m_e) should also vary in time. However, current uncertainty in R_U measurement still leaves a room for other alternative ideas and possible coincidences. For example, noting that $m_p/m_e \sim 40/3 \times \alpha^{-1}$, relation (1) can have another form:

$$R_U = \frac{m_p}{m_e}\frac{1}{4}\left(\frac{3}{10}\frac{ke^2}{m_e c^2}\right) 2^{128} \tag{1.3}$$

which would correspond to 13.95809 Gyr. As this value is currently out of the present WMAP data frame, therefore it is not supported by the author here.

2 Electron-proton radius coincidence

Another interesting idea connects the classical proton radius and gravitational radius of the electron by an exact equation as follows:

$$\frac{r_p}{r_{ge}} = 2^{128}, \tag{2}$$

where $r_p = \frac{1}{2}\frac{3}{5}ke^2/(m_p c^2)$ — classical proton radius and $r_{gp} = 2Gm_e/c^2$ — gravitational electron radius (i.e. the Schwarzschild radius for the electron mass). Of course some comments are required regarding coefficients $\frac{1}{2}$ and $\frac{3}{5}$. Usually numerical factors are ignored or assumed to be close to unity when defining classical (electron) radius. However, suggested new definition has exact numerical factor $\frac{3}{10} = \frac{3}{5} \times \frac{1}{2}$, so it is obvious to have the following explanations for that one by one:

- Ratio $\frac{3}{5}$ is classical proton radius definition. The only important difference with modern representation of the classical radius is the coefficient $\frac{3}{5}$. It is well known from electrostatics that the energy required to assemble a sphere of constant charge density of radius r and charge r is $E = \frac{3}{5}ke^2/r$. Usually these factors like $\frac{3}{5}$ or $\frac{1}{2}$ are ignored while defining the classical electron radius. Surprisingly, the coincidence advices the use of $\frac{3}{5}$ which means that charge is equally spread within the sphere of the electron (or proton) radius.
- Ratio $\frac{1}{2}$ in classical proton radius definition. Usual definition of the classical radius does not require having $\frac{1}{2}$ because initially one relates total electrostatic energy (E_e) of the electron (or proton) to rest mass energy as following: $E_e = mc^2$. The factor $\frac{1}{2}$ appears if one postulates that electromagnetic energy (E_{em}) of the electron or proton is just a half of particle's rest mass energy as: $E_{em} = \frac{1}{2}mc^2$. There are two possible alternative explanations for this:
  1. The Virial Theorem that tells us that the potential energy inside a given volume is balanced by the kinetic energy of matter and equals to half of it. So if one considers electromagnetic energy as kinetic and rest mass as potential energy we would have: $E_{em} = \frac{1}{2}mc^2$;
  2. Simply assuming that half of total energy may be magnetic energy or of another nature. One may also propose that there could be no $\frac{1}{2}$ in classical proton radius definition, but there is 2^{129} instead of 2^{129} in formula (2). From the author's point of view this does not correspond to reality, and particularly the number 2^{128} should have strong presence in all numerical expressions of Nature.

It can be easily seen that $r_p = (m_e/m_p)r_e$, so another way to rewrite (2) is:

$$\frac{r_e}{r_{ge}} = \frac{m_p}{m_e}\, 2^{128}. \tag{2.1}$$

And this leads to another possible representation of the initial formula as:

$$\frac{r_e}{r_{gp}} = 2^{128}, \tag{2.2}$$

where r_e is classical electron radius, r_{gp} is gravitational (Schwarzschild) radius of the proton. The expression (2.2) is very similar to (2). So, we may actually combine them into another interesting equation:

$$r_p\, r_{gp} = r_e\, r_{ge}. \tag{2.2a}$$

The precision of the Electron-proton coincidence given by (2) is smaller than 0.02%. From the author's point of view this deviation originates from current uncertainty in gravitational constant (G) measurement. If we consider that the relative G uncertainty nowadays is around and not less than 0.02% then we must accept this amazing and unexplained coincidence that allows us to predict the exact value for the gravitational constant (G). So, this finding suggests that the following possible consequences are valid. Firstly, because of $\frac{3}{5}$ ratio proton or electron still may be considered as classical particle with uniform charge density inside its radius. And secondly, directly from (2) one can express the value of the Newtonian constant of gravitation (G) exactly as follows:

$$G = \frac{3}{20}\frac{ke^2}{m_p m_e}\, 2^{-128}. \tag{2.3}$$

It leads to exact value for $G = 6.674632 \times 10^{-11}$. This value is within the frame of 2010 CODATA-recommended

value with standard uncertainty given by: $6.67384 \pm 0.00080 \times 10^{-11}$ [6] (See also figure). One may compare this expression with the similar one obtained in 1929 by R. Fürth [10]:

$$G = \frac{hc}{\pi(m_p + m_e)^2} = 16^{32}$$

that is read in SI units for G as:

$$G = \frac{2\hbar c}{(m_p + m_e)^2}\, 16^{-32}.$$

It is interesting to compare it to (2.3) to note obvious similarity. However, one may see that the expression is not satisfactory because it leads to the value of ($G = 6.63466 \times 10^{-11}$) which has significant deviation (0.59%) and is far out of 2010 CODATA range. So, the expression 2.3 (which fits well to modern data) is quite challenging because it may be confirmed or disproved by future experimental data for G.

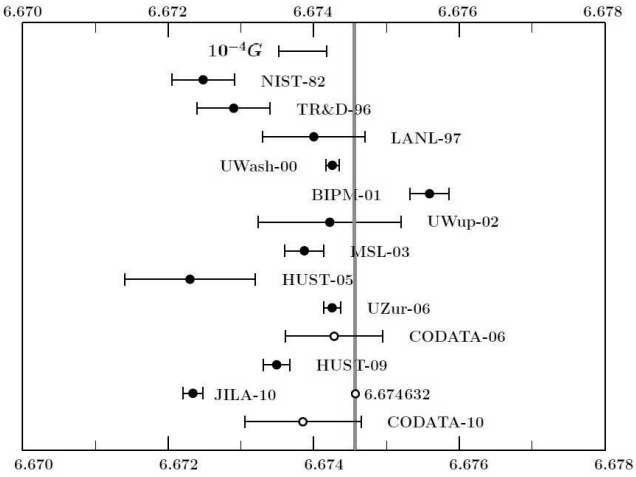

Fig. 1: The figure with recent experimental data for the Newtonian constant of gravitation G. The vertical line corresponds to the value obtained from (2.3).

Though the obtained value fits quite well into current experimental data, the author does not exclude some possible small deviations caused by vacuum polarisation and consequent slight deviation from the uniformity of the charge distribution (like Uehling Potential). So we will look at this in future works.

It is also important to stress that the use of classical proton radius here is very provisional and in principle could be avoided: so the same result for G may be obtained using only the electron's classical radius.

It is easy to note also that Dirac's Large Number N precisely equals to:

$$N = \frac{ke^2}{m_p m_e} = \frac{20}{3}\, 2^{128}. \tag{2.4}$$

This means that variation of Dirac's Large Number (N) in time is hardly possible, because 2^{128} represents simply the constant number. So the ratio of the electromagnetic force to the gravitational one remains always constant during the current epoch.

3 Weak cosmological coincidence

$$\frac{2c^3}{G} \approx \frac{m_p}{t_p}\, 2^{128}, \tag{3}$$

where c is speed of light, G is the gravitational constant, $t_p = \hbar/(m_p c^2)$-period of reduced Compton wave of the proton. This equation may be interpreted as relation of rate of mass growth or the expansion rate of the Universe [4, 5] to harmonic properties of the proton as wave. However the relative precision of (3) is 0.48% (or even 0.49% if we accept definition of G as in 2.3) which is unsatisfactory for modern measurements and it makes the expression valid only approximately. In order to become more precise the expression should have the following representation:

$$\frac{2c^3}{G} \approx \frac{m_p + 9m_e}{t_p}\, 2^{128}. \tag{3.1}$$

Or alternatively (to become exactly precise):

$$\frac{2c^3}{G} = \frac{m_p}{t_p}\frac{20}{3}\, \alpha^{-1}\, 2^{128}. \tag{3.2}$$

But further discussion of this topic will be explored in future works.

4 Weak electron-proton mass ratio

The attempts to explain large numbers by placing inverted fine structure constant in exponential function have been done many times before [11, 12]. Another interesting hypothesis could relate proton to electron mass ratio with fine structure constant and the number 2^{128} in the following manner:

$$\frac{m_p}{m_e} \simeq \frac{7}{2}\, 2^{(\alpha^{-1} - 128)}. \tag{4}$$

However the relative precision is comparably high (0.06%) and is out of the error frame of the current experimental data. However, using this relation as approximation, one can find similar connections of derived formulas to the similar ones in work [12].

Conclusion

The basic meaning of all these relations may be viewed in the form of exact equality for large Dirac's number N (see 2.4). However, all these proposals disprove one of the Dirac's hypothesis regarding the equality of the big numbers [2, see p. 200]. So, the author has shown that the number N, which is the ratio of the electromagnetic force to its gravitational force given by (2.4), is actually not equal to number N_1 which is

the ratio of Universe radius to classical electron radius (1.2). However these two differ only by the numerical factor of $20.55 = \alpha^{-1} \times (3/20)$. So, the main conclusions of this study are as follows:

- Current Universe age and radius can be calculated exactly (13.8896 Gyr);
- The value of Newtonian constant of gravitation (G) can be derived exactly (6.674632×10^{-11});
- The number 2^{128} should have a real significance in the constants of Nature.

Generally the concept of "power of two" could be regarded as having two properties in science. Firstly, it is digital (logical) math where power of two has common use. So this may support an idea of holographic concept of the Universe or some of the fractal theories. Secondly, it is used in wave mechanics, and it could be viewed in accordance with wave properties of the elementary particles in quantum physics. In terms of wave concept, the number of 2^{128} corresponds to the tone of 128-th octave or to some higher harmonic ("overtone") of the main tone. It is interesting to mention that a very close idea has been brought few years ago. The idea relates particles mass levels within two sequences that descend in geometric progression from the Plank Mass. Sub-levels are arranged in subsequence of common ratio which uses a power of 2 [7, 8]. The author is also very supportive to the point of view given in [13–15], however it is important to stress that physics should be free from approximate relations and should have only precise equalities and formulas. Some of the exact formulas which may help to support such general ideas have been presented in this work. If new suggested relations for Large Numbers are correct then it should probably lead to new search for its hidden meaning. As always, we must accept the fact that in often cases new findings lead to new questions instead of the answers and that might become a new challenge for new investigations and theories. Assuming that at least one of the discovered relations is correct in the future we may become a bit closer to the true view on physical reality.

Submitted on December 17, 2012 /Accepted on December 20, 2012

References

1. Eddington A.S. The Philosophy of Physical Science, Tarner Lectures, 1938, Cambridge At The University Press, 1939.
2. Barrow, J.D. The Constants of Nature From Alpha to Omega — The Numbers that Encode the Deepest Secrets of the Universe. Pantheon Books, 2002.
3. NASA website NASA's Hubble Finds Most Distant Galaxy Candidate Ever Seen in Universe, News release from NASA: Jan. 26, 2011, Release: 11-025.
4. Kostro L. Physical interpretation of the coefficients c/G, c^2/G, c^3/G, c^4/G, c^5/G that appear in the equations of general relativity.
5. Schiller C. General relativity and cosmology derived from principle of maximum power or force, 2008, arXiv: physics/0607090.
6. Peter J. Mohr, Barry N. Taylor, and David B. Newell CODATA Recommended Values of the Fundamental Physical Constants: 2010. March 15, 2012. arXiv: physics.atom-ph/1203.5425.
7. Riley B.F. Particle mass levels, 2009. viXra: 0909.0006.
8. Riley B.F. A unified model of particle mass. arXiv: physics/0306098.
9. Jarosik, N., et.al. Seven-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Sky Maps, Systematic Errors, and Basic Results, 2011, ApJS, 192, 14. arXiv: astro-ph.CO/1001.4744.
10. Fürth R. Versuch einer quantentheoretischen Berechnung der Massen von Proton und Elektron, Physikalische Zeitschrift, 1929, from Helge Kragh Magic Number: A Partial History of the Fine-Structure Constant, DOI 10.1007/s00407-002-0065-7, *Archive for History of Exact Sciences*, 2003, v. 57, 395–431.
11. Brandmüller J. and Rüchardt E. Die Sommerfeldsche Feinstrukturkonstante und das Problem der spektroskopischen Einheiten, *Die Naturwissenschaften*, 1950, v. 37, from R. Baggiolini On a remarkable relation between atomic and universal constants, *American Journal of Physics*, 1957, v. 25, 324.
12. Sirag, Saul-Paul Physical constants as cosmological constraints, *International Journal of Theoretical Physics*, 1983, v. 22(12), 1067–1089.
13. Sanchez F.M., Kotov V., Bizouard C. Evidence for a steady-state, holographic, tachyonic and super-symmetric cosmology, *Galilean Electrodynamics*, 2009, v. 20, Special Issues, no. 3, 43–53.
14. Sanchez F.M., Kotov V., Bizouard C. Towards a synthesis of two cosmologies, *Journal of Cosmology*, v.17, 2011.
15. Sanchez F.M., Kotov V., Bizouard C. Towards Coherent Cosmology, *Galilean Electrodynamics*, 2012.
16. Gamov G. Numerology fo the constants of nature. *Procedings of the National Academy of Science U.S.A.*, 1968, v. 59(2), 313–318.
17. Whitrow G.J. An Elementary Deviation of Eddington Number 10^{78}, *Nature*, 1952, no. 4302, 611–612.

Proper Space Kinematics

Sean Wade
P.O. Box 246, Highmount, NY, 12441. E-mail: seanwadePSK@verizon.net

It is desirable to understand the movement of both matter and energy in the universe based upon fundamental principles of space and time. Time dilation and length contraction are features of Special Relativity derived from the observed constancy of the speed of light. Quantum Mechanics asserts that motion in the universe is probabilistic and not deterministic. While the practicality of these dissimilar theories is well established through widespread application inconsistencies in their marriage persist, marring their utility, and preventing their full expression. After identifying an error in perspective the current theories are tested by modifying logical assumptions to eliminate paradoxical contradictions. Analysis of simultaneous frames of reference leads to a new formulation of space and time that predicts the motion of both kinds of particles. *Proper Space* is a real, three-dimensional space clocked by proper time that is undergoing a densification at the rate of c. Coordinate transformations to a familiar *object space* and a mathematical *stationary space* clarify the counterintuitive aspects of Special Relativity. These symmetries demonstrate that within the local universe stationary observers are a forbidden frame of reference; all is in motion. In lieu of Quantum Mechanics and Uncertainty the use of the imaginary number i is restricted for application to the labeling of mass as either material or immaterial. This material phase difference accounts for both the perceived constant velocity of light and its apparent statistical nature. The application of Proper Space Kinematics will advance more accurate representations of microscopic, macroscopic, and cosmological processes and serve as a foundation for further study and reflection thereafter leading to greater insight.

1 Introduction

The planets dancing in the heavens, an apple falling to earth each kindle curiosity about the dynamical universe. The mysteries of the unseen world and its apparent influences on daily life inspire wonder and imagination. Such observations drive the search for hidden constraints that govern the actions of atomic particles and molecules, ballistic objects, and celestial bodies. Guided by tools of logic, intuition, and creativity philosophers, scientists, and mathematicians strive to model laws that describe movement in each realm. Many years of disparate effort and the resulting accumulation of knowledge demonstrate that there are underlying commonalities that apply across all physical scales. This connectedness prompts the realization that searching for unifying first principles based upon fundamental aspects of space and time is an attainable goal. Understanding the foundation that the universe is built upon enables the continuing pursuit of deeper and more profound truths and further illuminates the miracle of human existence.

In 1905 Albert Einstein published his landmark work *Zur Elektrodynamik bewegter Körper* [1] (translated as *On the Electrodynamics of Moving Bodies* [2]). He stated that it was well known that under transformation to a moving reference frame Max-well's equations acquired asymmetries that were not present in nature. Einstein resolved these inconsistencies by introducing two fundamental principles [2]:

1. *The laws by which the states of physical systems undergo change are not affected, whether these changes of state be referred to the one or the other of two systems of co-ordinates in uniform translatory motion.*
2. *Any ray of light moves in the "stationary" system of co-ordinates with the determined velocity c, whether the ray be emitted by a stationary or by a moving body.*

The first postulate identified inertial frames of reference. The second postulate emphasized the constancy of the speed of light. From these followed the development of Special Relativity as a basis for motion.

Although the efficacy of Special Relativity cannot be denied it is a mathematical physics derived from the observations of light approaching any observer at the same speed regardless of the specific frame of reference. Any element of a theory that behaves identically under all applications must itself lie outside this theory and for this reason the action of discrete quanta requires a separate and distinctly different explanation.

This leads to the hard-fought and hard-won triumph of the Copenhagen interpretation of Quantum Mechanics culminating in its emergence as the preeminent theory of modern physics [3]. Owing to their experimental origins the composition of each theory contains mathematical elements that are not immediately obvious and consequently can act as obstacles to understanding and usage. If the basic realities of space and time are known then it is possible to properly explain the curious details of motion of all objects in the native environment and show that they proceed in a logical and intuitive way from this physical foundation.

This research is motivated by a personal failure of understanding attributable to the lack of a fundamental mechanics capable of explaining all rudimentary motion in the universe as derived from the basic condition of spacetime. Guided by instinct and curiosity the contemporary scientific theories and the corresponding philosophies are explored through a careful analysis of perspective; long-held premises are tested and discarded by virtue of the necessity to eliminate contradiction. The result of the methodology described in this paper addresses a kinematics which describes free motion without forces and interactions and with no regard for collisions and the associated quantities of energy, momentum and mass. A first principles theory is significant in that it can immeasurably improve physics on every level by serving as a foundation for the advancement of larger fields of research. The sluggish pace of grand unification, the overwrought complexity of string theory, the extremes of quantum gravity, the perplexity of dark matter, and the simplistic seeming three body problem are currently unresolved issues in physics [4]. These problems along with technological improvements to solar cell efficiency and medical scanning devices are among those that can potentially benefit from the application of Proper Space Kinematics.

2 Methodology

As a part of natural skepticism and scientific inquiry it is often useful to be able to replicate the research process both as a test of results and as a guide to understanding. In theoretical work much of the effort is introspective and it is impossible to retrace the labyrinthine mental pathways that lead to these results. In light of this difficulty it is practical to detail the initial impetus that motivated the author and to provide an overview of the techniques employed in the striving for enlightenment.

It is always more difficult to understand the fundamental principles that govern a system when the only perspectives available lie within the system itself. For this reason it is desirable to find a vantage point or frame of reference that lies outside the system so as not to be influenced by or subject to whatever constraints are imposed upon its occupants. In reviewing the basic elements of Special Relativity it is troubling that there are inconsistencies in the currently used theory between the common explanations and the mathematical model. While the equations purport to explain motion from an exterior viewpoint it is a theory of *relative* motion that performs as if a massive object occupies the choice of origin. This fallacy compounds the suspicion that an accurate picture of reality may not be known and necessitates the need for further exploration of this phenomena the source of which must thereafter be inferred from these confused aspects. In a similar mien the self-circular reasoning involved in using light itself as a mediator to measure lightspeed is also an obstacle to understanding and conceals basic mechanisms that are vital to accurately model the system mathematically.

Other concerns arise from a thoughtful analysis of the present philosophy. If the lightspeed barrier is a limiting condition then this implies that the velocity of an object is a more important kinematical consideration than position or acceleration. A cursory examination of the invariant interval suggests that its spatial and temporal components act in opposition to each other across varied reference frames although the use of hyperbolic functions would conversely imply a conjunction of underlying influences. The question of balance imparts an impression of rotation along a spectrum instead of a deviation from zero which is compounded by the inability to rotate a vector of zero length and might lead to the conclusion that nothing is static. The Quantum Mechanical proposition that the universe is unknowable at its most basic level and the ensuing enigma of wave-particle duality raise further reservations. Intuitively the structure of the universe should be based on the least number and simplest of principles although wisdom dictates that allowances be made for the possibility of deliberate design.

Logic is a weak tool for dissecting a system that is known to have defects in its application and for this reason a trial-by-solution is likely to be ineffective. Therefore the course of action must include an exploration using physical intuition and not only a mathematical manipulation of equations. This is accomplished through repeated testing of both implicit and explicit assumptions to find the origins of paradoxical situations and then to remove these faults. The movements of both energy and matter in spacetime are studied with careful consideration of perspective in an attempt to unravel the knot of relativity and to imagine an extrauniversal viewpoint. Producing an accurate answer to the dilemmas detailed here requires substantial time for trial calculations, for searching through potentially applicable literature, and for reevaluating conventional concepts through quiet reflection.

3 Results

The natural universe is undergoing a process of *densification* and is described here as being composed of three real spatial coordinates and one real monodirectional temporal counter. Densification is defined for this demesne as an increase in the density of space that occurs in the measure of distance between any two disparate points clocked by proper time and progressing at the rate of c. Previously referred to as lightspeed the particular value of the *characteristic velocity* as it has been measured serves as a label for the universe as well as all residents. It is further assumed that the inhabited universe is infinite though possibly bounded, is fixed relative to any preternatural background, if one exists, and is not undergoing additional physical alteration. The kinematics of finitesimal objects is derived for the movement of noninteracting rigid bodies traveling at constant speed. The premise of constant speed translates across all *spaces*. Initially this derivation is done without the qualification of particles as either matter or

energy. For the purposes of this paper it is practical and sufficient for understanding to consider equations of motion of only one dimension since any path traversed at constant speed can be parameterized as such and densifies at the same rate; extrapolation to all three dimensions is a straightforward task.

Length and time are measured with a ruler and a clock [5]. *Proper Space* is denoted by the variable z and experiences densification dependent on proper time which is denoted by the independent variable tau τ. In this case the clock is also embedded within the ruler and is not considered an additional physical dimension. In *object space* space and time are treated on equal footing as independent dimensions and are denoted by x and t, respectively. These variables have local values that manifest densification as contraction and dilation in mimicry of many of the details of Special Relativity and continue to suffer from dependence on frame-specific relative velocity.

Measurements of physical observables are made in object space and converted to values in proper space where the action originates. The coordinate transformation for length or displacement involves the scale change

$$dz = f\,dx. \tag{1}$$

The unitless scale factor f is defined for densification as a *density of points* which is represented by a ratio of infinities increasing from unity as

$$f = \left(\frac{dx + cdt}{dx}\right) = 1 + \left(\frac{cdt}{dx}\right). \tag{2}$$

Simple substitution of (2) into (1) yields the coordinate transformation between spaces

$$dz = dx + cdt. \tag{3}$$

This is the conversion for points in space with an explicit dependence on elapsed time. Contrary to expectation with densification a scale transformation from object space to proper space takes a form that is reminiscent of a Galilean boost [6].

The burgeoning density of proper space requires the use of additional notation for the proper *waxing velocity*, denoted by w, while in object space the concept of velocity is retained as it is traditionally used and remains denoted by v. The relationship between the two quantities is

$$w \triangleq \frac{dz}{d\tau} = \alpha\,(v + c)\,. \tag{4}$$

Values for the velocity in object space persist within the range of $(-c, c)$ while values for the waxing velocity are always positive within the range of $[c, 0)$. Open endpoints of each interval are forbidden for the same reason; denizens of the universe must always experience the advancement of proper time in some nonzero fraction. Accordingly values for the *temporal dilation coefficient*, marked by alpha α, vary as $[1, 0)$. Infinite dilation is taboo and is expressed by the avoidance of an asymptotic value of zero for α.

In a break from prior theories of motion it is important that velocities in all spaces are measured from a special class of perspectives hereinafter referred to as *proper frames*. The choice of axes may be made without particular regard for position but must be boosted to the specific velocity whereby t reaches the maximum expression of τ and experiences densification at its fullest flowering. Proper frames can be thought of as critical points and specific values associated with these perspectives are $w = c$, $v = 0$ and $\alpha = 1$.

For the sake of completeness it is worthwhile to also define a *stationary space*, denoted by y, which advances with the preceding variable of proper time τ. This nonphysical construct may be mathematically advantageous as it allows for the use of global variables that forgo dependence on relative perspective but carries the caveat that the space is not demonstrative of physical reality. The scale-densification — to — boost technique above is repeated to provide the transformation to proper space as

$$dz = dy + cd\tau. \tag{5}$$

Measurements of length or distance are converted from object space to corresponding values in stationary space through the transitive property with application of (3) and (5) to yield

$$dy + cd\tau = dx + cdt. \tag{6}$$

For stationary space a *pseudovelocity* is defined as u and takes on the values $(-c, c)$. Values of u are somewhat analogous to velocities v in object space e.g., adopting the value of zero in a proper frame where $dt = d\tau$. The relation for the two quantities is

$$u \triangleq \frac{dy}{d\tau} = \alpha\,(v + c) - c. \tag{7}$$

As proper space and stationary space both share the variable τ as proper time the relationship between velocities is more simple as

$$w = u + c. \tag{8}$$

The choice of alphabetically proximate variables is a mnemonic convenience that is intended to be familiar and resemble current definitions but not to imply any other mathematical relationship including equivalence with commonly used spatial unit vectors. The invariant variable s is reserved for possible future use.

4 Discussion

Change is the true nature of the universe and the densification of proper space depicts the most authentic representation of space and time. A static ruler of fixed length is a forbidden item; an absolutely stationary observer is a nonsensical frame of reference that does not exist. Although this picture of reality is not mathematically convenient it is the correct philosophy to accurately model basic kinematics. Object(-*ive*) space is the milieu where action is perceived and measurements are made. The coordinate transformation to proper

space takes the form of a boost centered on c which arises, not surprisingly, from the defining feature of the universe. This conversion yields the advantage to the waxing velocity which can always be rotated since it is never zero as objects must experience some positive slice of proper time. Objects moving at the same rate as densification do not experience the passage of proper time and therefore cannot inhabit this universe. It should not be overlooked that the transformation is originally a scale change whereby the size of massive objects is growing relative to the coordinate system with the densification. It is the growth of the span between the center of mass of an object and any other contained point within that same object that is seemingly retarded in entities not occupying a proper frame of motion. Consideration of the action of only infinitesimal points does not reveal this detail. It is helpful if the time-dependent metric tensor is visualized as the ruler growing shorter and shorter thus creating an illusion of inflation. The author supposes that the idea of densification within fixed boundaries is an option that Einstein either discarded or failed to consider and is the source of his self-critical vacillation regarding the Cosmological Constant [7].

In a brief departure from kinematics an examination of multiple perspectives clarifies the necessity for a preferred frame of interaction. Collisions cannot have different outcomes in different frames otherwise every incident can be transformed into a destructive event. Synchronization to a proper frame is a sufficient condition to preserve the integrity of any physical interaction; the regimentation also reemphasizes the significance of velocity. This interpretation of simultaneity provides the means to intellectually resolve the well-known gedanken paradox [8]: what are the ages of the traveling twins? There currently exists an abundance of experimental and observational data which can be used to determine the validity of proper frames. The incongruity of superluminal travel can be rectified by application of the results discussed here and the presence of tachyons is discarded.

Terminology relating to motion must be used cautiously since the concepts involved vary among the different spaces despite a similarity in formulation. Calculations done in stationary space remove some of the difficulties of perspective that are inherent to the other spaces but readers are warned to remember that this is not a physical reality. In object space it is time that slows and space that contracts as a function of speed to the detriment of the occupying objects. A sequence of snapshots in proper space shows that movement in any direction produces an apparent spatial *and* temporal dilation based upon the movement of a mass impinging on the budding densification. Part of the virtue of proper space is that the object itself is not actually altered and the perception of dilation occurs only in the direction of motion while densification continues unabated along all other axes. Along with the increase in movement this retardation of proper space and proper time is demonstrated as a decrease in the waxing velocity although the moving particle still perceives densification continuing at c. A reasonable choice for a functional definition of w is the hyperbolic secant as a function of the *angle of dilation*, represented by phi φ, and demonstrated in $w = \operatorname{csech}(\varphi)$ making it more akin to a speed than a velocity. The positive-definite, even function is a rotation of phi through the real interval $(-\infty, \infty)$ as measured from a proper frame and this run equates with the previously detailed bounds for w of $[c, 0)$. The choice of hyperbolic functions is preferred over the circular transcendental equations as the hyperbolics are independent of the imaginary number i.

Consideration of the relative velocity between bulk objects with determinate length requires the use of a proper frame. A measurement of relative velocity is inadequate to completely determine the true states of objects in the system; two measurements are required to establish the correct scalings for space and time. Take the example in object space of two masses *at rest* to a specific proper frame as well as to each other; the waxing velocity of each frame in proper space is c. While the relative velocity in object space between the centers of mass remains at zero in proper space the relative velocity is characteristic and not zero as might be anticipated. This discrepancy can be partly reconciled by acknowledging the supplemental velocity acquired in proper space which is imparted by the densification of the gap between the two masses. Accordingly the correct velocities between the center of mass frames are emphasized by primed coordinates and subscripts enumerate the frames of reference for separate and distinct objects as

$$u' \triangleq u_2 - u_1, \tag{9.1}$$

$$w' \triangleq u' + c, \tag{9.2}$$

$$\text{and } v' \triangleq \left(\frac{w'}{\alpha'}\right) - c. \tag{9.3}$$

The sense of relative motion is preserved by these transformations; the distinction of an *alias* versus an *alibi* transformation is highlighted [9]. To determine the relative velocity in object space measurements are made there first, converted to pseudovelocities and the relative velocity calculated then reverted to object space. All direct measurements are relative with v' equal to v from a proper frame. Although this computation avoids direct expression of quantities in proper space the kernel of the action lies there.

The primed alpha coefficient α' serves as both the relative temporal dilation between objects as well as the transformation between frames in proper space. It is defined as a ratio in the range of real positive numbers $(0, \infty)$ and is most easily understand as an exponential with argument given as the difference between two angles and shown here

$$\alpha' \triangleq \left(\frac{\alpha_2}{\alpha_1}\right) = e^{-(\varphi_2 - \varphi_1)}. \tag{10}$$

These definitions in combination with some computation restore the hyperbolic tangent in a composition of velocities in

object space and yield a result that is in correspondence with rapidity [10]. The assertion that values of α' can exceed 1 is a specific deficiency in the conventional measurement of relative velocity. Attend to these calculations with care as variables of the traditional theory are ill-defined by the muddled use of mixed perspective due to a misconception in the choice of laboratory frame.

The derivation is accomplished to this point without the need for i; further descriptions of the manifest complexity of nature require the use of imaginary numbers. The kinematics is extended to distinguish between the movements of the two types of mass by applying a label of *material* or *immaterial* (i-material) to all particles whether they are matter or energy; the two types are interchangeable provided the exchange is done en masse. Real and imaginary objects occupy overlapping worlds within the same universe because the phase dichotomy causes a perception of near invisibility between the two categories of mass in which the contrary object collapses to a dimensionless point. As seen before with dilation the flattening is perceptual and not actual. The alternately phased object appears to ignore densification and to therefore exist in a forbidden state. In that the object doesn't seem to experience scaling it performs as with a waxing velocity of zero and erroneously claims relative velocities as $w' = c \pm c$ and $v' = 0 \pm c$. The relative motion of the oppositely phased objects either approaches or recedes depending on the relative angle of dilation. The tipping point occurs when $\varphi_1 = \varphi_2$ and $\alpha' = 1$ and can serve as a test provided it is possible to produce a series of identical immaterial objects. The author defers the specific method for this production to the expertise of experimentalists.

The expression $E = mc^2$ acquires a new complexion after revisiting the outmoded concepts of the rest mass of matter and the mass equivalence of energy. The characteristic velocity measured between real and imaginary particles is superficial and acts as a screening value whereby information is hidden from the casual observer but still preserved. Relying only on light as a mediator to comprehend motion introduces inaccuracies that must be corrected. A single physical measurement of an immaterial object is underdetermined and wrongly constrains the associated parameters of velocity and imaginary mass. Consequently the sources of wave nature are found to originate from the complex quality of mass and not directly from the tableau of spacetime. The seeming lack of determinate states which is the hallmark of Quantum Mechanics illustrates its subservience to statistical models and elucidates its failure of completeness and its misappropriation of fundamental reality.

5 Conclusions

Maintaining an open-minded attitude of skepticism lies at the heart of the scientific method; challenging established ideas is not necessarily an effort towards rebellion and anarchy. Persistent testing is an important undertaking in the quest to further humanity's understanding of life, the universe and everything. The author is awakened to the fact that the peculiar consequences of Albert Einstein's Special Relativity and subsequent geometric interpretation of space and time originate from observation and the theory does not proceed directly from a foundational source. Relying on relative viewpoints to predict motion has an inherent handicap and in combination with the confused measurement of lightspeed initially serve as motivation for study. The approach to creating a kinematics involves keeping a critical eye on perspective and attempting to dispel paradoxes in order to see through to the metaphysical center. It is a mistake to rely totally on mathematical models of nature as they are ultimately flawed and physicists must constantly endeavor to look beyond constructed images of reality. If the basic realities of space and time are known then it is possible to properly explain the curious details of motion of all objects in the native environment and show that they proceed in a logical and intuitive way from this physical foundation. The success of such a hypothesis would be the pedestal on which the future of physics could be built and would have a far-reaching influence on science and greatly impact its application to technology in addition to answering important philosophical questions.

The elegance of Proper Space Kinematics is that it proceeds directly from the fundamental concept that the fabric of the universe densifies at the unique quantity and quality of the characteristic velocity c maturing with an inescapable duration of proper time. This insight into the inner workings of space and time solidifies realizations regarding the arrow of time and the spectre of irreversible entropy. It is not surprising that in a study of motion appearances are deceiving and this deception necessitates a transformation to positions in other spaces which are difficult to visualize since the use of a time-dependent metric is not a well-developed field of study with much pertinent literature. Spatial densification is understood by a study of the steadily mounting density of points (Mind the infinities!) whereby a scale change converts the growing size of objects to the form of a boost. Care at the beginning: reconceptions of velocity and movement lead to new definitions such as proper space's waxing velocity and the interrelated temporal dilation coefficient. Additionally boosting perspective to any proper frame provides the linkage that shows these points of view can be logically related and provides for surety over the use of four-vectors and four-velocities. Scrutiny of these results discerns that stationary space is a fictitious point of view that proves to be a useful tool.

Densification clarifies the observed nuances of motion more clearly than Special Relativity by eschewing stationary states and shedding new light on the evolution of the aging universe. Scale expansion of objects is found to be a new source of motion where movement hinders the passage of time and limits experience. Thought problems are revisited

and explained by the introduction of new concepts such as proper frames providing ample opportunity for testing the validity of these new ideas; experimental and mathematical verification have many available avenues to explore. This kinematics shows that the movement of objects does not cause a physical change but merely alters appearances. As particles always experience their own perspective as characteristic the presented composition of velocities accurately details the difference between spectators and participants. The duality of mass shows that the landscape of space is a perpendicular reality for matter and energy which can be tested by manufacturing an experimental watershed. The screening between imaginary material phase shifts creates a Quantum confusion due to underdetermined measurements that the author feels does finally vindicate Einstein's intuition. (No Dice!)

As a first principles theory which meets the onus of the stated hypothesis Proper Space Kinematics claims jurisdiction over all motion in the universe. Proper motion supplants the golden relics of relative and absolute motion; the dubious lessons of Quantum nature must be extracted and distilled for their essential truths. As seen with Isaac Newton in his 1687 *Philosophiae Naturalis Principia Mathematica* [11] in the continuing quest for deeper insight new ideas are a harbinger for chaos as fundamental changes in understanding prompt the reevaluation of physics on every level and in every niche. The potential impact on science and its application expands from the theoretical to the technological to hopefully improve the quality of human life and reinvigorate the search for profundity. The author proposes that the next step in this study is to complete a mechanics in full generality with metric-tensor formalism to include a derivation of canonical coordinates with energy and momentum and an examination of accelerating objects with interactions via both collisions and forces-at-a-distance. Delving further raises a rich multitude of questions: Is densification in the universe constant? What does this mean for cosmology and the birth and death of the universe? Are there other characteristic parallel universes that are unseen? Is there a greater realm? How do these results apply to the standard model? Was the creation of life and homo sapiens sapiens an accident? Why are we here? Physicists have always searched the universe for bedrock on which to stand but to live in harmony with our world we must instead navigate the rising tide of space and time and learn to walk on water.

Acknowledgements

Although this research has been a solitary exercise I wish to express my gratitude to those who have lifted me up and helped make me capable of achieving this goal. To my family and friends, first and foremost of which are my parents Paul A. Wade and Margaret Aracich Wade, thank you for your love and support in helping me to build a place within myself where I can always stand strong. Thank you Dr. Stephen J. Nettel for your role in awakening my love of physics. Thank you to all of those people who have cared about and for my well-being for without each of you this work would never have been realized.

Submitted on December 21, 2012 / Accepted on December 24, 2012

References

1. Einstein A. Zur Elektrodynamik bewegter Körper. *Annalen der Physik*, 1905, Bd. 17, 891–921.
2. Einstein A. On the electrodynamics of moving bodies. In: *The Principle of Relativity* (Methuen and Company, Ltd., London, 1923), translation by W. Perrett and G. B. Jeffery, from *Das Relativatsprinzip* (Tuebner, Germany, 1922), 4th ed.
3. Faye J. Copenhagen Interpretation of Quantum Mechanics. *The Stanford Encyclopedia of Philosophy*, 2008, http://plato.stanford.edu/archives/fall2008/entries/qm-copenhagen
4. List of unsolved problems in physics. http://en.wikipedia.org/wiki/List_of_unsolved_problems_in_physics
5. Schouten J. A. Tensor Analysis for Physicists. Dover Publications, Inc., Mineola, 1989, 2nd ed., p. 217.
6. Misner C. W., Thorne K. S., and Wheeler J. A. Gravitation. W. H. Freeman and Company, New York, 1973, p. 295.
7. Mook D. E. and Vargish T. Inside Relativity. Princeton University Press, Princeton, 1987, p. 191.
8. Hartle J. B. Gravity: An Introduction to Einstein's General Relativity. Addison Wesley, San Francisco, 2003, p. 63.
9. Weisstein E. W. Transformation. http://mathworld.wolfram.com/Transformation.html
10. Foster J. and Nightingale J. D. A Short Course in General Relativity. Springer Science & Business Media, Inc., New York, 2006, p. 216.
11. Newton I. Philosophiae Naturalis Principia Mathematica. Jussu Societatis Regiae ac Typis Josephi Streater, Londini, 1687.

Liquid Metallic Hydrogen II. A Critical Assessment of Current and Primordial Helium Levels in the Sun

Pierre-Marie Robitaille
Department of Radiology, The Ohio State University, 395 W. 12th Ave, Columbus, Ohio 43210, USA. E-mail: robitaille.1@osu.edu

Before a solar model becomes viable in astrophysics, one must consider how the elemental constitution of the Sun was ascertained, especially relative to its principle components: hydrogen and helium. Liquid metallic hydrogen has been proposed as a solar structural material for models based on condensed matter (e.g. Robitaille P.-M. Liquid Metallic Hydrogen: A Building Block for the Liquid Sun. *Progr. Phys.*, 2011, v. 3, 60–74). There can be little doubt that hydrogen plays a dominant role in the universe and in the stars; the massive abundance of hydrogen in the Sun was established long ago. Today, it can be demonstrated that the near isointense nature of the Sun's Balmer lines provides strong confirmatory evidence for a distinct solar surface. The situation relative to helium remains less conclusive. Still, helium occupies a prominent role in astronomy, both as an element associated with cosmology and as a byproduct of nuclear energy generation, though its abundances within the Sun cannot be reliably estimated using theoretical approaches. With respect to the determination of helium levels, the element remains spectroscopically silent at the level of the photosphere. While helium can be monitored with ease in the chromosphere and the prominences of the corona using spectroscopic methods, these measures are highly variable and responsive to elevated solar activity and nuclear fragmentation. Direct assays of the solar winds are currently viewed as incapable of providing definitive information regarding solar helium abundances. As a result, insight relative to helium remains strictly based on theoretical estimates which couple helioseismological approaches to metrics derived from solar models. Despite their "state of the art" nature, helium estimates based on solar models and helioseismology are suspect on several fronts, including their reliance on solar opacities. The best knowledge can only come from the solar winds which, though highly variable, provide a wealth of data. Evaluations of primordial helium levels based on 1) the spectroscopic study of H-II regions and 2) microwave anisotropy data, remain highly questionable. Current helium levels, both within the stars (Robitaille J. C. and Robitaille P.-M. Liquid Metallic Hydrogen III. Intercalation and Lattice Exclusion versus Gravitational Settling, and Their Consequences Relative to Internal Structure, Surface Activity, and Solar Winds in the Sun. *Progr. Phys.*, 2013, v. 2, in press) and the universe at large, appear to be overstated. A careful consideration of available observational data suggests that helium abundances are considerably lower than currently believed.

> At the age of five Cecilia [Payne] saw a meteor, and thereupon decided to become an Astronomer. She remarked that she must begin quickly, in case there should be no research left when she grew up.
>
> Betty Grierson Leaf, 1923 [1, p. 72–73]

1 Introduction

Knowledge that helium [2, 3] was first observed in the Sun by Pierre Jules César Janssen [4] and Joseph Norman Lockyer [5], before being discovered on Earth by William Ramsay [6], might prompt the belief that the element was abundant on the solar surface. In fact, helium has never been identified in the absorption spectra of the quiet Sun. Janssen and Lockyer's fortunate discovery was restricted to helium lines appearing within the prominences of the corona and within the disturbed chromosphere [4, 5]. While the element was easily detectable in these regions [7], helium has remained relatively spectroscopically silent on the Sun. Conversely, the stars and the Sun display signs of extreme hydrogen abundance, as first observed by Cecilia Payne [8], Albrecht Unsöld [9], and Henry Norris Russell [10]. Few would take issue with the conclusion that the visible universe is primarily comprised of hydrogen. Helium abundances present a more arduous question.

Despite all the difficulties, several lines of reasoning sustain the tremendous attention that solar helium levels have received in astronomy. First, helium is the end product of the nuclear reactions currently believed to fuel many of the stars, either in the pp process or the CNO cycle [11–15]. Second, solar helium levels are inherently linked to the gaseous models of the Sun [16–18] and the application of theoretical findings to the interpretation of helioseismic results [19–23]. Finally, helium is thought to be a key primordial element in

Big Bang cosmology [3, 24–30]. As a result, the evaluation of helium levels in the Sun brings a unified vision of astrophysics, wherein accepted solar values lend credence to our current concept of the formation of the universe. Still, questions remain relative to the accuracy of modern helium determinations.

A flurry of initial studies had suggested that helium abundances in the stars approached 27% by mass (see [3] for a review). The findings provided support for those who proposed primordial formation of helium prior to the existence of the objects which populate the main sequence [3, 24]. However, these ideas were challenged when it was discovered that certain B-type stars, which should have been rich in helium lines, were almost devoid of such features [3]. As a result, in certain stars, helium was said to be gravitationally settling towards the interior [3, 31]. The desire to link helium levels in the Sun with those anticipated from the primordial synthesis continues to dominate modern solar theory [18]. Nonetheless, it can be demonstrated that the methods used to estimate primordial helium levels in the universe [24] are either highly suspect or implausible. Given these complexities, it is appropriate to compose a critical review of how helium abundances have been historically obtained and how they are currently determined, both in the Sun and in the universe at large.

2 Assessing elemental abundances in stellar spectra

2.1 The Saha Equations

Reasoning, like Lindemann [32] and Eggert [33] before him, that the fragmentation of an atom into an ion and an electron was analogous to the dissociation of a molecule, Megh Nad Saha [34, 35] formulated the ionization equations [36, 37] in the early 1920s. In so doing, he called upon the Nernst equation [38] and suggested that the free electron could be viewed as an ideal gas. He also relied on thermal equilibrium and the ionization potentials of the elements. Since Saha's equation was inherently related to parameters associated with the ideal gas (i.e. [39, p. 29–36] and [40, p. 107–117]) he demonstrated that the level of ionization could be increased either with elevated temperature or decreased pressure. Saha hypothesized that the pressure of the reversing layer approached 0.1–1 atm [36, p. 481] and was the first to utilize this assumption to account for the appearance of spectral lines across stellar classes as simple functions of temperature [36, 37]. He was concerned with the marginal appearance of spectral lines [36, 37], that point at which these features first appeared on a photographic plate. Cecilia Payne [1, 41] would soon estimate the abundance of the elements in the universe using the same criterion [8].

In his initial work, Saha would comment on the impossibility of solar temperatures increasing as one moves from the photosphere to the upper chromosphere: "*Lockyer's theory... [that elements become more ionized as higher elevations are reached within the chromosphere] ... would lead us to the hypothesis that the outer chromosphere is at a substantially higher temperature than the photosphere, and the lower chromosphere; and that the temperature of the sun increases as we pass radially outwards. This hypothesis is, however, quite untenable and is in flagrant contradiction to all accepted theories of physics*" [36, p. 473]. Saha had not suspected that 20th century solar theorists would maintain such a position. Lockyer's analysis was correct: ionization increased with elevation in the chromosphere. This was an important lesson relative to thermal equilibrium. In any case, Saha did observe that hydrogen was not fully ionized in the chromosphere, since the lines from H_α and H_β were evident at this level. He also recognized that hydrogen should be essentially ionized in O class stars and that the lines coincident with the Balmer series in these stars had originated from ionized helium. At the same time, he outlined that the same spectral lines for classes later than B2A were completely due to hydrogen [37, p. 151].

Subrahmanyan Chandrasekhar's (Nobel Prize, 1983 [42]) thesis advisor, Sir Ralph H. Fowler [43], had provided significant insight and criticisms into Saha's second manuscript [37, p. 153] and the resulting text was masterful. In 1927, Megh Nad Saha was elected a Fellow of the Royal Society [34].

In the meantime, Fowler [43] and Edward Arthur Milne [44] would collaborate and construct a wonderful extension [45, 46] of Saha's seminal papers [36, 37]. They improved the treatment of ionization to consider not only principle lines arising from atoms in their lowest energy states, but also the subordinate lines produced by excited atoms and ions [45, 46]. For his part, Saha had concentrated on the excitation and ionization of the neutral atom [36, 37]. Fowler and Milne understood that the marginal appearance of a spectral line could be used in determining relative concentrations and provided some indication of the minimum number of atoms necessary for appearance [45, 46]. They emphasized the idea that: "*the intensity of a given absorption line in a stellar spectrum is proportional to the concentration of atoms in the stellar atmosphere capable of absorbing the line*" [45, p. 404]. Their first paper also highlighted the value of the maximum of a spectral line in assessing the temperature and pressure of the reversing layer and outlined that this problem was not affected by the relative abundance of the element studied [45]. Using stellar data from the lines of Ca, Mg, Sr, and Ba they determined that the electron pressure of the reversing layer was on the order of 10^{-4} atm [45]. Fowler and Milne understood that electron pressure, P_e, of the reversing layer was not determined by a single ionization process, but by the ionization of many elements: "*In thus regarding P_e as fundamental we are in effect assuming that, due to the presence of more easily ionised atoms, there are so many electrons present that the partial electron pressure is practically independent of the degree of ionization of the element under discussion*" [45, p. 409]. They expressed concern that their results led to the assumption that absorbing species had very large absorption

coefficients [45]. Milne had already determined that the absorption coefficients should be very large [47] and would later devote another theoretical paper to their determination [48]. In their work together, Fowler and Milne explicitly assumed that the reversing layer could be treated as existing under conditions of thermal equilibrium, as Saha's treatment required [36]. The validity of such assumptions is not simple to ascertain.

At Cambridge, Milne met Cecilia Payne [1, p. 121], a student at Newnham College [1, p. 112] and learned of her impending access to the vast collection of photographic plates used to generate the Henry Draper Catalogue at the Harvard Observatory [1, p. 144–153]. Prior to the advent of the modern MKK classification [49], the Henry Draper Catalogue was the largest stellar library collection, with over 200,000 classified stars [1, p. 144–153]. Milne suggested that "*if he had... [Payne's] ... opportunity, he would go after the observations that would test and verify the Saha theory*" [1, p. 155]. Cecila Payne soon left Cambridge and sailed to America.

2.2 Cecilia Payne: What is the universe made of?

> "I remember when, as a student at Cambridge, I decided I wanted to be an astronomer and asked the advice of Colonel Stratton, he replied, "You can't expect to be anything but an amateur". I should have been discouraged, but I wasn't, so I asked Eddington the same question. He (as was his way) thought it over a very long time and finally said: "I can see no insuperable obstacle" [50, xv].

Nineteenth century scientists had little on which to base their understanding of the composition of the universe. Their clues could only come from the Earth itself and from the meteorites which occasionally tumbled onto its surface. Consequently, it was not unreasonable to expect that the universe's composition matched the terrestrial setting. However, stellar spectra, already stored on photographic plates throughout Europe and especially in the vast Henry Draper Collection, were hiding a drastically altered viewpoint. With the arrival of yet another woman at the Harvard Observatory [51–60], the stars could not much longer conceal their story. Surrounded by Pickering's Harem [51–60], Cecilia Payne [1, 41] completed her classic report on the abundance of the elements [8] and became the first to underscore the importance of hydrogen as the constitutive atom of universe. Her thesis had been carefully prepared and presented supportive laboratory evidence, not only of ionization potentials, but of the validity of Saha's treatment [8, p. 105–115].

Stellar spectra signaled hydrogen [61] was so abundant that several scientists, including Henry Norris Russell, could not fully accept the conclusion. Payne had written an early manuscript detailing the tremendous presence of hydrogen [1, p. 19]. Her thesis advisor, Harlow Shapley, forwarded the work to Russell who commented: "*It is clearly impossible that hydrogen should be a million times more abundant than the metals*" [1, p. 19]. That early manuscript was never published and has since been lost [1, p. 20]. Tempered by Russell and Shapley, Cecilia Payne finally produced her famous PhD dissertation: *Stellar Atmospheres: A Contribution to the Observational Study of High Temperature in the Reversing Layers of Stars* [8]. She would comment on hydrogen in this manner: "*Although hydrogen and helium are manifestly very abundant in stellar atmospheres, the actual values derived from the estimates of marginal appearance are regarded as spurious*" [8, p. 186]. A little later she would add: "*The outstanding discrepancies between the astrophysical and terrestrial abundances are displayed for hydrogen and helium. The enormous abundance derived for these elements in the stellar atmospheres is almost certainly not real*" [8, p. 188] and "*The lines of both atoms appear to be far more persistent, at high and low temperatures, than those of any other element*" [8, p. 189].

For her part, Payne privately maintained that hydrogen was tremendously abundant in the stars: "*When I returned to visit Cambridge after I finished this first essay in astrophysics, I went to see Eddington. In a burst of youthful enthusiasm, I told him that I believed that there was far more hydrogen in the stars than any other atom. 'You don't mean in the stars, you mean on the stars', was his comment. In this case, indeed, I was in the right, and in later years he was to recognize it too*" [1, p. 165].

Payne's work also highlighted the importance of helium in the O and B class stars [8]. For the first time, hydrogen and helium became the focus of scrutiny for their role as potential building blocks of the stars and the cosmos [8]. She emphasized that: "*there is no reason to assume a sensible departure from uniform composition for members of the normal sequence*" [8, p. 179] and "*The uniformity of composition of stellar atmospheres is an established fact*" [8, p. 189]. She also held, as Eddington and Zeipel had advanced, that given their gaseous nature: "*an effect of rotation of a star will be to keep the constituents well mixed, so that the outer portions of the sun or of a star are probably fairly representative of the interior*" [8, p. 185]. Still, Payne was cautious relative to extending her results as reflecting the internal composition of the stars: "*The observations on abundances refer merely to the stellar atmosphere, and it is not possible to arrive in this way at conclusions as to internal composition. But marked differences of internal composition from star to star might be expected to affect the atmosphere to a noticeable extent, and it is therefore somewhat unlikely that such differences do occur*" [8, p. 189].

Payne would conclude her thesis with a wonderful exposition of the Henry Draper Classification system [8, p. 190–198]. Otto Struve would come to regard the study as "*the most brilliant Ph.D. thesis ever written in astronomy*" [41]. Edwin Hubble would comment relative to Payne: "*She's the best man at Harvard*" [1, p. 184]. As Milne suggested, the first dissertation of the Harvard College Observatory was founded

upon the application of the ionization equations [36,37,45,46] to the detailed analysis of spectral lines across stellar classes. It did not specifically address elemental abundances in the Sun. Nonetheless, Payne's 1925 dissertation heralded the application of quantitative spectral analysis in astronomy [8].

2.3 Albrecht Unsöld, hydrogen abundance, and evidence for a solar surface

Albrecht Unsöld extended Payne's studies with a focus on the solar spectra [9]. Following in her footsteps [8], in 1928 [9], he applied the ionization formula [36, 37] to the chromosphere and estimated the levels of sodium, aluminum, calcium, strontium, and barium. In addition, Unsöld determined that the electron gas pressure in the chromosphere stood at $\sim 10^{-6}$ atm [9]. He also concluded that hydrogen must be about one million times more abundant than any other element in the Sun [9, 62]. William McCrea was soon to echo Unsöld, finding that hydrogen was a million times more abundant than Ca^{+} within the chromosphere [62, 63].

Importantly, Unsöld also documented that the absorbance of the hydrogen β, γ, and δ lines did not decrease across the Balmer series ($H_\alpha = 1$; $H_\beta = 0.73$; $H_\gamma = 0.91$; $H_\delta = 1.0$) as expected from quantum mechanical considerations ($H_\alpha = 1$; $H_\beta = 0.19$; $H_\gamma = 0.07$; $H_\delta = 0.03$) [9]. This was an important finding relative to the nature of the Sun. Recently, the behavior of hydrogen emission lines has been analyzed with non-LTE methods [64]. It has been concluded that the "*$n = 3$ and higher levels are in detailed balance deep in the photosphere, but they develop a non-LTE underpopulation further out. However, the levels with higher n-values stay in detailed balance relative to each other at these atmospheric depths, and they also collisionally couple tightly to the continuum*" [64]. Yet, in the gaseous models of the Sun, the continuum is not composed of condensed matter [65]. It represents an area of profoundly increased solar opacity [65]. Nevertheless, the behavior of the Balmer series in the solar atmosphere strongly supports the idea that the Sun is comprised of condensed matter. Only a physical entity of sufficient density, such as a surface, can permit tight collisional coupling to the continuum, as it is impossible to couple to the opacity changes which characterize the continuum in gaseous models [65]. These findings comprise the sixteenth and seventeenth lines of evidence that the Sun is comprised of condensed matter. The others are outlined by the author in recent publications (e.g. [66]).

2.4 Henry Norris Russell: Inability to estimate Helium from spectral lines

Soon Henry Norris Russell [67] surpassed Unsöld in his analysis of solar spectral lines and provided a detailed compositional analysis of the Sun. Relative to the occupied energy levels within atoms on the Sun, Russell affirmed that: "*It must further be born in mind that even at solar temperatures the great majority of the atoms of any given kind, whether ionized or neutral, will be in the state of lowest energy*" [10, p. 21]. At the same time, Russell realized that this rule was not observed by hydrogen, leading him to the conclusion that the element was extremely abundant in the Sun: "*One non-metal, however, presents a real and glaring exception to the general rule. The hydrogen lines of the Balmer series, and, as Babcock has recently shown, of the Paschen series as well, are very strong in the Sun, though the energy required to put an atom into condition to absorb these series is, respectively, 10.16 and 12.04 volts - higher than for any other solar absorption lines. The obvious explanation — that hydrogen is far more abundant than the other elements — appears to be the only one*" [10, p. 22]. In fact, even the hydrogen Brackett lines can be visualized in the infrared spectrum of the Sun [68]. Russell also highlighted Unsöld's observation [9] that the hydrogen β, γ, and δ lines did not decrease as expected. That the hydrogen lines were extremely broad in the Sun had already been well established. Russell echoed some of his contemporaries and suggested that this might result from a Stark effect [10, p. 50].

Finally, Russell accepted Payne's findings relative to hydrogen and reported her numbers for the elements without comment in his table XVI [10, p. 65]. He stated that: "*The most important previous determination of the abundance of the elements by astrophysical means is that by Miss Payne...*" [10, p. 64]. Russell found the correlation between their works to display "*a very gratifying agreement*" [10, p. 65]

Like Payne, Russell had relied on the work of Fowler and Milne [45, 46] to set the composition of the Sun. He implemented their suggestion that electron pressures, P_e, could be gathered by considering the spectra and the ionization potential for elements like Ca, Sc, Ti, Sr and Yt. From these, he deduced a P_e of 3.1×10^{-6} atm, in close agreement with Milne (2.5×10^{-5} atm), and Payne and Hogg (2.54×10^{-6} atm) in class G0 stars [10, p. 54–55]. Along with John Quincy Stewart, Russell had previously considered various means of determining the pressures at the Sun's surface and had determined that the pressure of the reversion layer could not be more than 10^{-4} atm [69]. But Russell reported a factor of at least 10 in discordance in calculating electron pressures based on either the ionization formula or the numbers of metallic atoms and ions [10, p. 70–71]. He would resolve the difficulty at the end of his treatise when setting the final elemental composition for the Sun [10, p. 72].

At the same time, while Payne had understood the importance of local thermal equilibrium (LTE) for the proper application of Saha's equation [8, p. 92–101], she did not attempt to make an explicit correction for the lack of equilibrium. Conversely, Russell placed a correction factor in his work for departure from LTE: "*We have finally to take into consideration the fact that the atmosphere may not be in thermodynamic equilibrium. The comparison of solar and stellar spectra affords evidence that this is the case*" [10, p. 52]. Relative

to his final abundances he commented: "*The main source of uncertainty which affects them is the magnitude of the correction for departure from thermodynamic equilibrium*" [10, p. 58] and "*If the correction for departure from thermodynamic equilibrium should be wholly disregarded, the calculated abundance of hydrogen — already very great — would be increased thirty fold*" [10, p. 62]. In the 1920s, of course, there was hesitancy concerning the tremendous levels of hydrogen observed in the solar atmosphere.

For Russell, oxygen appeared as abundant as all other metals combined. He also argued against, although did not fully dismiss, gravitational settling in the Sun for the heaviest metals: "*It does not appear necessary, therefore, to assume that downward diffusion depletes the sun's atmosphere of the heavier elements, though the possibility of such an influence remains*" [10, p. 59]. Importantly, he noted: "*The statement that enhanced lines are found in the sun for those elements which have lines of low excitation potential in the accessible region has therefore few exceptions*" [10, p. 35]. At the same time, he advanced that for those elements "*which fail to show enhancement lines in the sun, the excitation potentials for the accessible lines are high in every case for which they have been determined*" [10, p. 35]. Furthermore Russell hypothesized that: "*It appears, therefore, that the principle factor which is unfavourable to the appearance of a spectral line in the sun is a high excitation potential*" [10, p. 35]. This was precisely the case relative to helium.

With respect to the second element, Russell wrote: "*There is but one element known to exist in the sun for which no estimate of abundance has now been made - and this is He. The intensity of its lines in the chromosphere shows that it must be present in considerable amount, but no quantitative estimate seems possible*" [10, p. 62]. Here was an explicit admission that solar helium abundances could not be ascertained using spectral data.

Helium was abundantly visible in early type stars, as Cecilia Payne had already discovered [8] and Paul Rudnick [70] and Anne Underhill continued to confirm [71–73]. Estimates of the number of hydrogen to helium atoms in O and B type stars varied from values as low as 3.2 to more than 27 [73, p. 156]. A factor of nearly 10 in relative abundances from spectral lines in such stars was hardly reassuring. Nonetheless, Underhill still surmised that the number of helium atoms was at the 4–5% level [73]. Yet for the Sun, data about helium abundance remained wanting.

2.5 Local Thermal Equilibrium

Milne was perhaps the greatest authority relative to local thermal equilibrium (LTE) in astronomy [74–77] and many of the most salient aspects of his arguments have been reviewed [78]. Milne advocated that LTE existed in the center of a star and that his treatment permitted "*us to see in a general way why the state of local thermodynamic equilibrium in the interior of a star breaks down as we approach the surface*" [77, p. 81–83]. In 1928, Milne would express concern relative to the appropriateness of the inferred thermal equilibrium in the reversing layer, as required by the Saha equations [36, 37], although he believed that studies based on the validity of the ionization equations should be pursued: "*The recent work of Adams and Russell brings forward evidence that the reversing layers of stars are not in thermodynamic equilibrium. This suggests a degree of caution in applying the fundamental method and formulae of Saha to stellar spectra. Nevertheless, departure from thermodynamic equilibrium can only be found by pushing to as great a refinement as possible the theory which assumes thermodynamic equilibrium*" [48]. Gerasimovic had already advanced corrections for small deviations from thermal equilibrium [79] and Russell applied corrections directly in his work [10]. By 1925, the Saha equations had been generally confirmed under experimental conditions (e.g. [8, p. 111–112] and [80]), but only in the broadest sense. Over time, the ionization equations continued to be widely studied and the problems considered were extended to include two-temperature plasmas (e.g. [81]), high pressures (e.g. [82]), varying opacities (e.g. [83]), and non-LTE (e.g. [84–88]). The Saha equations eventually became a useful staple in the treatment of plasma physics [89, p. 164] and stellar atmospheres [90–92].

As Auer highlighted relative to solar models [88], under non-LTE, a set of rate equations enters into the problem of determining the abundance of any given electronic state. Furthermore, the radiation field is introduced directly into the equations [88] utilized to calculate both opacities and populations. The problem therefore becomes dependent on "*simultaneous knowledge of the radiation field at all frequencies and all depths*" [88, p. 576].

While ionization appeared tractable given modern computing, the solution became linked to the knowledge of stellar opacities, an area of theory whose weaknesses have already been outlined [78]. Nonetheless, non-LTE approaches have been successful in addressing the spectra of early type stars [93–95]. Today, such methods also account for electronic, atomic, and ionic collision processes [64]. Non-LTE approaches have provided considerable insight into the Balmer and Paschen series associated with the hydrogen spectrum of the Sun [64].

Finally, it appears that the treatment adopted by Cecilia Payne might not have been too far afield [8]. For many of the cooler stars, simple LTE seems sufficient to address ionization problems [94]. Non-LTE methods become most important for the O and A class stars [93–95]. In any case, helium cannot be assessed on the Sun using the ionization equations due to the lack of appropriate spectral lines. As a result, while the LTE and non-LTE settings may be fundamental to the proper treatment of spectral lines, the methods have little bearing on the proper evaluation of helium levels in the Sun.

3 Helium from solar theory

3.1 Henry Norris Russell

Since Russell was not able to extract helium abundances directly from spectral lines, he did so, without further scientific justification, by assuming that the Sun had an mean molecular weight of ∼ 2 [10, p. 72–73]. Such a value had also been suggested by Saha [36, p. 476], who had in turn adopted it from Eddington [96, p. 596]. As for Eddington, he had previously examined the radiation equilibrium of the stars using a mean molecular weight of 54 [97]. In 1916, this value had been selected based on the belief that the stars were principally composed of elements such as oxygen, silicon, and iron prior to full ionization [1, viii]. Eddington lowered the mean molecular weight to a value of 2 in 1917 [96, p. 596], based on the idea that the elements would be fully ionized in the stars. In the fully ionized state, hydrogen has a mean molecular weight of 0.5, helium of ∼ 1.3, and iron of ∼ 2 (see [40, p. 102–104] for a full discussion of mean molecular weights in astrophysics). It was this value which Russell was to adopt in his calculations.

Using a mean molecular weight corresponding to a metal rich star, Russell concluded that helium was 13% as abundant as hydrogen by weight [10, p. 73]. He then computed that the Sun had equal percentages of oxygen and other metals (∼ 24% each) and that hydrogen comprised just under half of the constitution (∼ 45%) by weight (see table XX in [10, p. 73]. If Russell had selected a mean atomic weight of ∼ 0.5, there would be dramatic changes in the calculated helium levels.

3.2 Early abundance calculations

In arbitrarily selecting mean molecular weights [96, 97], Eddington determined the mean central stellar temperatures and pressures along with the acceleration due to gravity at the surface (e.g. [97, p. 22]). In turn, these parameters altered the calculated absorption coefficient, and hence opacity, of stellar interiors [97, p. 22]. Consequently, the setting of mean atomic weight had a profound implication on nearly every aspect of stellar modeling, but opacity would always remain paramount. In 1922, Eddington had derived a relationship between opacity and temperature [98] which would become known as Kramer's law [99].

Soon, Strömgren introduced an interesting twist to Eddington's approach [100, 101]. Rather than assuming a mean atomic weight, Strömgren began his calculations by computing opacity values, and from there, estimating the fractional composition of hydrogen within several stars [100], relying in part on Russell's elemental composition [10]. He concluded that the fractional abundance of hydrogen was ∼ 0.3 and maintained that the presence of helium would have little effect on these calculations since "*hydrogen and helium do not contribute to the opacity directly*" [100, p. 139]. Strömgren would write: "*we have neglected the influence of helium. The helium proportion is rather uncertain and the error introduced by neglecting helium altogether small* [100, p. 142]. Modern stellar theory would come to rely greatly on the opacity contributions of the negative hydrogen ion (H^-) [102]. Strömgren's assumptions were premature. Still, he championed the idea of initially computing opacity, and from these values obtaining both solar parameters and elemental abundances [100, 101].

Following the publication of a key modeling paper by Cowling [103], Martin Schwarzschild was to take the next theoretical step [104]. First, he made use of the mass-luminosity relation while expressing mean molecular weight and opacity as a function of elemental composition (X = hydrogen, Y = helium) [104]. Then, reasoning that the energy output in the Sun from the CNO cycle [13] was directly related to elemental composition, he derived a fractional elemental composition for hydrogen, helium, and the metals equal to 0.47, 0.41, and 0.12, respectively [104]. The results were once again critically dependent on estimated opacities, which Schwartzchild, like Strömgren before him [100, 101], assumed to display Eddington's [98] −3.5 power dependence on temperature (see Eq. 9 in [104]). In fact, Schwarzschild utilized an even greater dependence on temperature for energy production, allowing a 17th power in the exponential (see Eq. 11 in [104]). Yugo Iinuma then advanced a broader approach to the stellar composition problem [105]. He was concerned with ranges of reasonable starting points, both for hydrogen concentration and average molecular weight. His treatment remained dependent on opacity computations, though less rigid in its conclusions [105]. Schwarzschild et al. [106] then introduced the effects of inhomogeneity in the solar interior and convective envelopes along with solar age into the abundance problem. They reached the conclusion that the temperatures at the core of the Sun were such that the carbon cycle should start to contribute to the problem. Hydrogen abundances were assumed in order to arrive both at a convection parameter and at helium values [106]. The critical link to opacity remained [106]. Weymann, who like Schwarzschild, was also at the Institute for Advanced Study, built on his findings [107]. Taking account of the carbon cycle, Weymann found that the core of the Sun was not convective [107]. Powers of 4 and 20 for temperature were assumed in the energy generation laws associated with the pp and CNO cycles [107]. The hydrogen fractional composition of the Sun was assumed and ranged from 0.60 to 0.80 (see Table 3 in [107]). This resulted in helium and metallic fractional compositions of 0.19–0.32 and 0.01–0.08, respectively (see Table 3 in [107]).

In 1961, Osterbrock and Rogerson would elegantly summarize the situation relative to estimating helium abundances in the Sun: "*Though helium is observed in the upper chromosphere and in prominences, the physical conditions in these regions are too complicated and imperfectly understood for the abundance ratio to be determined from measurements of*

these emission lines. Hence the only reliable way to find the helium abundance in the Sun is by analysis of its internal structure" [108]. Yet, given the progress to date, the determination of elemental compositions within the Sun had been a complex adventure involving either assumed values of average molecular weights, hydrogen abundances, energy generation reactions, and opacity. The latter would eventually present the greatest difficulties [78]. Osterbrock and Rogerson would utilize Weymann's calculation, along with making an assumption by setting the Z/X ratio at 6.4×10^{-2} [108], to estimate interior solar fractional abundances at $X = 0.67$, $Y = 0.29$, and $Z = 0.04$. They were guided in this estimation by the belief that: "*the solar, planetary nebula, and interstellar abundances are all essentially the same*" [108, p. 132]. For the planetary nebula NGC 7027 they set the fractional abundances at $X = 0.64$, $Y = 0.32$, and $Z = 0.04$ [108]. Solar elemental composition became decidedly linked to estimates from remote objects. The stage was set for conclusively linking solar elemental composition to stellar evolution and primordial nucleosynthesis.

3.3 Modern abundance calculation

Eventually, the solar neutrino problem entered theoretical modeling [16, 109]. In his simulations, John Bahcall would utilize fractional abundances of relatively narrow range ($X = 0.715 - 0.80$, $Y = 0.19 - 0.258$ and $Z = 0.01 - 0.027$), setting the central densities and temperatures near 150 g/cm^3 and 15 million Kelvin, respectively [16]. The results, as before, were reliant on the use of solar opacity estimates [78]. By the beginning of the 1970s, fractional abundances for helium and the metals were settling on values near 0.28 and 0.02 [25]. Solar models became increasingly complex, relying on stellar opacity tables [110–118], energy generation equations, neutrino flux, and solar age to arrive at internally consistent results [17, 18]. Complexity was also introduced by considering helium and heavy element diffusion throughout the solar body [17, 18, 119, 120]. It became important to establish not only modern helium content, but also the initial helium abundance in the Sun [17,21,121]. Gough had already suggested that helioseismology could be used to help establish fractional abundances: "*Thus one might anticipate inferring the hydrogen-helium abundance ratio by comparing the measured values with a sequence of model solar envelopes*" [19, p. 21]. Helioseismological results became strongly incorporated into solar modeling [20–23] and "*helioseismic techniques ... [became] ... the most accurate way to determine the solar helium abundance*" [20, p. 235]. The techniques remained linked to the equations of state which contained six unknowns including: elemental composition, density, temperature, and pressure [20, p. 224]. Moreover, the problems required an explicit knowledge of opacity [20, p. 224] from its associated tables [110–118].

Relative to solar models, the central problem remains linked to the determination of internal solar opacity. The questions are complex and have been addressed in detail already by the author [78]. In the end, opacity tables [110–118] have no place in the treatment of stellar problems, precisely because they are incapable of reproducing the thermal emission spectrum required [78]. They simply mask ignorance of a fundamental problem in astronomy: the mechanism for the production of a thermal spectrum. Their inability to account for the production of a single photon by graphite on Earth [78], establishes that stellar opacity derived from isolated atoms and ions can play no role in the proper understanding of thermal emissivity in the stars. As a result, helium levels can never be established using theoretical modeling based on the gaseous equations of state and their inherent association with stellar opacity tables [78].

4 Primordial helium abundances

The quest to understand helium levels in the stars has been further complicated by the inferred association of this element with primordial nucleosynthesis in Big Bang cosmology [24–30]. Early on, Alpher, Bethe, and Gamow postulated that the elements had been synthesized in a primordial fireball [122]. This nucleosynthesis was proposed to include the entire periodic table and even unstable elements, with short lifetimes, of greater atomic number [122]. Soon, the idea that the composition of the stars was largely related to primordial conditions was born, especially relative to hydrogen and helium [24, 123]. No other scheme appeared likely to explain the tremendous He levels in stellar atmospheres, which approached 27% by weight [3, 24]: "*It is the purpose of this article to suggest that mild 'cooking' [such as found in stars] is not enough and that most, if not all, of the material of our everyday world, of the Sun, of the stars in our Galaxy and probably of the whole local group of galaxies, if not the whole Universe, has been 'cooked' to a temperature in excess of* $10^{10}K$" [123, p. 1108]. By then, the astrophysical community had already accepted that the heavy elements, which constituted trivial amounts of matter compared to hydrogen and helium, had largely been synthesized in the stars [14]. Only ^{1}H, ^{2}H, ^{3}He, ^{4}He, and ^{7}Li became candidates for synthesis through a primordial process [124, 125].

The postulate that "*helium abundance is universal and was generated in a Big Bang*" [125] eventually came to wide acceptance. The entire theory was hinged on elevated helium abundances: "*We can now say that if the Universe originated in a singular way the He/H ratio cannot be less than about 0.14. This value is of the same order of magnitude as the observed ratios although it is somewhat larger than most of them. However, if it can be established empirically that the ratio is appreciably less than this in any astronomical object in which diffussive seperation is out of the question, we can assert that the Universe did not have a singular origin*" [123, p. 1109]. Elevated helium levels, along with the discovery

of the microwave background [126] and the red-shifts of distant galaxies [127, 128] became one of the three great pillars of Big Bang cosmology [24, 129, 130]. This explained why gravitational settling had become critical in discounting low helium abundances of certain B type stars [3, 30, 31]. If empirical helium levels fell into question and a mechanism existed to accept the tremendously decreased helium levels in these special B type stars [3, 31] by preventing gravitational settling [131], Big Bang cosmology could not survive. Stellar and solar helium abundances cannot be allowed to drop in modern cosmology.

Today, the quest to link helium abundances and primordial nucleosynthesis has continued [26–30] using two lines of reasoning: 1) the analysis of anisotropy in the microwave background [132, 133] and 2) the observation of helium and hydrogen lines from low-metallicity extragalactic HII regions [26, 134–137].

Unfortunately, the use of anisotropy data [132, 133] to analyze primordial helium abundances are highly suspect. First, insurmountable problems exist with the WMAP data sets, as already highlighted by the author [138]. WMAP suffers from significant galactic foreground contamination which cannot be properly removed [138]. In addition, the WMAP team cannot distinguish between signal arising from a hypothetically primordial origin from those produced throughout the universe as a result of normal stellar activity [138]. While evident 'point sources' are taken into account, it remains impossible to determine, on a pixel by pixel basis, whether the signal has a primordial origin, or originates from an unidentified non-cosmological object [138]. Furthermore, WMAP raw data has proven to be unstable from year to year in a manner inconsistent with the hypothesized cosmological origins of these signals [138]. The data suffers from poor signal to noise and the ILC coefficients used for generating the final anisotropy maps do not remain constant between data releases [138]. Most troubling, the data sets cannot be combined using a unique combination of spectral channels [138]. As a result, since no unique anisotropy data set can be extracted [138], the data has no scientific value in analyzing helium abundances. Similar problems will occur when data from the Planck satellite finally becomes available [139]. As a result, all helium abundances derived from microwave anisotropy data sets must be viewed with a high degree of suspicion.

On the surface, the extraction of primordial helium abundances from H II regions appears more feasible [26, 134–137]. H II regions are rich in both hydrogen and helium but have low heavy element abundances (~ 1/40 solar) [140]. Unlike H I regions (~ 60K), H II regions exist at temperatures between 7,500 and 13,000 K [141]. In H II regions "*the* ^{4}He *abundance is derived from the recombination lines of singly and doubly ionized* ^{4}He*; neutral* ^{4}He *is unobserved*" [140, p. 50]. Unfortunately, experiments which utilized H II regions to assess primordial helium cannot easily ascertain that the sample has a uniform elemental composition. Furthermore, the use of H II regions for this purpose discounts the idea that helium has been synthesized locally. Such a suggestion should not be easily dismissed, as the temperatures of observation [141] are well above those in equilibrium with the hypothesized residual temperature of the Big Bang (~ 3K) [130]. Only low metallicity supports the idea that these helium concentrations are primordial. Nothing should prevent stellar systems from creating regions of low metallicity outside of a cosmological context. In this regard, the elevated temperatures of H II regions suggest that a process well beyond primordial considerations is now influencing elemental abundances in these regions. As such, it is imprudent to derive primordial helium abundances from H II regions.

We do not know, and will probably never be able to ascertain, primordial helium abundances. In order to observe helium in astronomy, elevated temperatures are required. These immediately imply that the processes observed are no longer in thermal equilibrium with those of interest in cosmology [130].

5 Solar winds: The key to understanding helium

Helium abundances can also be monitored in the solar wind [143–152]. Presumably, the results are so dynamic that they cannot be utilized to establish helium levels in the Sun itself. However, solar winds [143–152] have presented astronomy with a wealth of scientific information, which could be used to profoundly alter our understanding of the Sun [131].

Already in 1971, it was recognized that solar wind helium abundance measurements gave values which were lower than those ascertained from theoretical experiments [143, p. 369]. The study of solar winds became linked to models of the corona. Although the relative abundance and velocities of hydrogen to helium were advanced as profoundly dependent on location [143], it remained evident that solar winds harbored a great deal of reliable information. Early on, it was known that helium to hydrogen density ratios in the solar wind could experience dramatic fluctuations [144], especially in slow winds [147], though values appeared more stable at high solar wind speeds [145]. Extremely low ratios of 0.01, rising to 0.08, with an average of 0.037, were reported [144]. Clearly, such values were in direct conflict with the elevated helium levels expected in the Sun from primordial arguments [123]. As such, solar wind measurements became viewed as unreliable relative to estimating helium abundances in the Sun [148].

Nonetheless, something truly fascinating was present in solar wind data. The Sun appeared to be expelling helium (J. C. Robitaille, personal communication [131]) with increased activity. The helium to hydrogen ratio was observed to increase in association with the onset of geomagnetic storms [144] and was highly responsive to the solar cycle [146, 149, 151]. The helium abundance could rise from average values of less than 2% at the solar minimum to around 4.5% at maximum [149]. After the early 1970s, the vari-

ation in solar wind helium abundance became increasingly pronounced. By 1982, helium abundances in the solar wind came to vary from values as low as 0.001 to as elevated as 0.35 [147]. A single value as high as 0.40 was reported [147]. At least half of all elevated helium abundance events were related to a transient interplanetary shock wave disturbance [147], though a significant portion were not associated with such events. Each of these extremes highlighted something phenomenal relative to solar winds. To explain the variability, theoretical models turned to the large scale structure of plasma. It was assumed that elevated helium abundance originated in regions of high magnetic field activity in the corona [131]. It was found that helium abundance "*enhancements often have unusually high ionization temperatures, indicative of an origin in active solar processes... Collectively, these observations suggest that... [helium abundance] ... enhancements in the solar wind signal the arrival of plasma ejected from low in the corona during a disturbance such as a large solar flare or an eruptive prominence*" [147]. While solar winds had a close link to the "*composition of the source material*" it could then "*be modified by the processes which operate in the transition zone and in the inner corona*" [148]. Primordial helium abundances within the Sun could be saved by discounting that solar wind helium abundances had any meaning whatsoever relative to the composition of the Sun itself. The idea that solar activity reflected the expulsion of helium from the Sun (J. C. Robitaille, personal communication [131]) was never advanced. While the scientific community maintained that helium abundances were not reliable, they claimed that it was possible to ascertain the fractional isotopic composition of the elements in the solar wind and relate them directly to the solar convective zone: "*The variability of the elemental abundances in the solar wind on all time scales and the FIP... [first ionization potential] ... effect, and its variability, will make it difficult to derive accurate solar abundances from solar wind measurements, with the exception of isotopic determinations*" [150]. Of course, isotope analysis could never constitute a challenge to the existence of large amounts of primordial helium in the Sun [123]. Solar wind helium abundances had to be simply correlated to the coronal magnetic field, although the correlation coefficient was not powerful ($\sigma \sim 0.3$) [152]. Nonetheless, helium abundance depressions could not be explained under such a scenario [152]. At the same time, it is currently believed that "*solar wind abundances are not a genuine, unbiased sample of solar abundances, but they are fractionated. One such fractionation depends on the first ionization potential (FIP): When comparing solar wind to solar abundances, elements with low FIP (<10 eV) are enriched by a significant factor, the FIP bias, over those with a high FIP ... Another fractionation process affects mainly helium, causing its abundance in the SW to be only about half of the solar abundance... It is most likely due to insufficient Coulomb drag between protons and alpha particles in the accelerating solar wind*" [154, p. 16]. Herein was an explicit admission that the cause of extremely low helium levels in the solar wind could not be adequately understood. Conversely, fractionation models continued to insist that elevated helium abundances were linked to the fractionation of large atoms by collisions with protons [152, 153]. Nothing could be gathered about solar helium abundances from solar winds precisely because theoretical constructs forbade such conclusions.

6 Conclusions

Modern day reports of elemental abundances in the Sun [154–156] maintain that the Sun has a relatively large proportion of helium with Y values typically near 0.248 and primordial values of 0.275. These values come from theoretical modeling, as helium remains spectroscopically silent in the photosphere and solar winds are viewed as unreliable [155, p. 166]. Therefore, claims that helium has "*very high abundance*" [155, p. 166] in the Sun are not supported by observational fact. In the end, mankind understands much less about this central element than a cursory review of the literature might suggest. Careful consideration of solar modeling establishes that all theoretical estimates of helium levels in the Sun cannot be relied upon, given their dependence of solar opacity tables [78]. This also applies to theoretical results which attempt to extract helium levels from helioseismology [156]. For this reason, it is simply not possible to establish elevated helium levels in the Sun from theory. As helium levels cannot be established spectroscopically, we are left with the solar winds for guidance.

Currently, solar winds are viewed as too complex to yield information relative to solar abundances. In large measure, this is because scientists are trying to understand this data in the context of an object whose helium abundance has been largely set in primordial times [24, 123, 155]. The idea that the Sun and the stars are actively working to control their helium levels has never been previously considered [131]. Nevertheless, the association of solar activity and elevated helium levels [146, 149, 151] strongly suggests that the active Sun is expelling helium and excluding it from its hydrogen based lattice (J. C. Robitaille, personal communication [131]). Herein can be found the cause of extremely low helium abundance often obtained in the slow solar wind: the Sun works to keep its helium levels low and solar activity represent a direct manisfestation of this fact. In the quiet Sun the slow solar winds can report fractional abundances of less than 2% and these should be viewed as steady state helium removal from the convective zone of the Sun. Such an idea strongly supports the contention that the Sun and the stars are primarily comprised of hydrogen in the liquid metallic state [131, 157].

In advancing that the universe is largely composed of hydrogen and that helium is being excluded from the stars (J. C. Robitaille, personal communication [131]), perhaps it is appropriate to turn once again to Cecilia Payne, as the first as-

tronomer to highlight the tremendous abundance of hydrogen in the universe [8]. As a child, she had been eager to become an astronomer "*in case there should be no research left when she grew up*" [1, p. 72–73]. Yet, her position changed dramatically with age: "*Looking back on my years of research, I don't like to dwell only on my mistakes; I am inclined to count my blessings, and two seem to me to be very especially valuable. The first blessing is that the process of discovery is gradual — if we were confronted with all the facts at once we should be so bewildered that we should not know how to interpret them. The second blessing is that we are not immortal. I say this because, after all, the human mind is not pliable enough to adapt to the continual changes in scientific ideas and techniques. I suspect there are still many astronomers who are working on problems, and with equipment, that are many years out of date. Now that I am old, I see that it is dangerous to be in too much of a hurry, to be too anxious to see the final result oneself. Our research does not belong to us, to our institution, or to our country. It belongs to mankind. And so I say to you, the young generation of astronomers: more power to you. May you continue to expand the picture of the universe, and may you never lose the thrill it gave you when it first broke on you in all its glory*" [Cecilia Payne-Gaposchkin, April 10, 1968 [50, p. xv]].

Dedication

This work is dedicated to my oldest son, Jacob.

Submitted on: December 25, 2012 / Accepted on: December 31, 2012
First published in online on: January 6, 2012

References

1. Haramundanis K. Cecilia Payne-Gaposchkin: An autobiography and other recollections (2nd Edition). Cambridge University Press, Cambridge, U.K., 1996.
2. Kragh H. The solar element: A reconsideration of helium's early history. *Ann. Science*, 2009, v. 66(2), 157–182.
3. Taylor R.J. The helium problem. *Quart. J. Roy. Astron. Soc.*, 1967, v. 8, 313–333.
4. Janssen J. Indications de quelques-uns des résultats obténus à Guntoor, pendant l'éclipse du mois d'août dernier. *Comptes Rendus*, 1868, v. 67, 838–39.
5. Lockyer J.N. Notice of an observation of the spectrum of a solar prominence. *Proc. Roy. Soc. London*, 1868, v. 17, 91–92.
6. Ramsey W. On a gas showing the spectrum of helium, the reputed cause of D3, one of the lines in the coronal spectrum. Preliminary note. *Proc. Roy. Soc. London*, 1895, v. 58, 347–352.
7. Hirayama T. The abundance of helium in the prominences and in the chromosphere. *Solar Physics*, 1971, v. 19, 384–400.
8. Payne C.H. The relative abundances of the elements. Stellar Atmospheres. Harvard Observatory Monograph no. 1 (Harlow Shapley, Editor), Harvard University Press, Cambridge, MA, 1925 (reprinted in part in Lang K.R. and Gingerich O. A source book in astronomy and astrophysics, 1900–1975, Harvard University Press, Cambridge, MA, 1979, p.245–248).
9. Unsöld A. Über die Struktur der Fraunhofersehen Linien und die quantitative Spektralanalyse der Sonnenatmosphäre. *Zeitschrift für Physik*, 1928, v.46(11–12), 765–781.
10. Russell H.N. On the composition of the Sun's atmosphere. *Astrophys. J.*, 1929, v.70, 11–82.
11. Bethe H.A. and Critchfield C.L. The formation of deuterons by proton combination. *Phys. Rev.*, 1938, v. 54, 248–254.
12. Bethe H.A. Energy production in stars. *Phys. Rev.*, 1939, v. 55(1), 103.
13. Bethe H.A. Energy production in stars. *Phys. Rev.*, 1939, v. 55(1), 434–456.
14. Burbidge E.M., Burbidge G.R., Fowler W.A. and Hoyle F. Synthesis of the elements in the stars. *Rev. Modern Phys.*, 1957, v. 29(4), 547–650.
15. Wallerstein G., Icko I., Parker P., Boesgaard A.M., Hale G.M., Champagne A.E., Barnes C.A., Käppeler F., Smith V.V., Hoffman R.D., Timmes F.X., Sneden C., Boyd R.N., Meyer B.S. and Lambert D.L. Synthesis of the elements in stars: forty years of progress. *Rev. Modern Phys.*, 1997, v. 69(4), 995–1084.
16. Bahcall J.N., Bahcall N.A. and Shaviv G. Present status of the theoretical predictions for the 37Cl solar-neutrino experiment. *Phys. Rev. Lett.*, 1968, v. 20(21), 1209–1212.
17. Bahcall J.N. and Pinsonneault M.H. Standard solar models, with and without helium diffusion, and the solar neutrino problem. *Rev. Mod. Phys.*, 1992, v. 64(4), 885–926.
18. Bachall J.N., Pinsonneault M.H. and Wasserburg G.J. Solar models with helium and heavy-element diffusion. *Rev. Mod. Phys.*, 1995, 67(4), 781–808.
19. Gough D.O. Towards a solar model. *Mem. Soc. Astron. Ital.*, 1984, v. 1–2, 13–35.
20. Basu S. and Antia H.M. Helioseismology and solar abundances. *Phys. Rep.*, 2008, v. 457, 217–283.
21. Serenelli A.M. and Basu S. Determining the initial helium abundance of the Sun. *Astrophys. J.*, 2010, v. 719(1), 865–872.
22. Delahaye F. and Pinsonneault M.H. The solar heavy-element abundances. I. Constraints from stellar interiors. *Astrophys. J.*, 2006, v. 649, 529–540.
23. van Saders J.L. and Pinsonneault M.H. The sensitivity of convection zone depth to stellar abundances: An absolute stellar abundance scale from asteroseismology. *Astrophys. J.*, 2012, v. 746, 16 (16 pages).
24. Peebles P.J.E. Primordial helium abundance and the primordial fireball. II. *Astrophys. J.*, 1966, v. 146, 542–552.
25. Danzinger I.J. The cosmic abundance of helium. *Ann. Rev. Astron. Astrophys.*, 1970, v. 8, 161–178.
26. Izotov Y.I. and Thuan T.X. The primordial abundance of 4He revisited. *Astrophys. J.*, 1998, v. 500, 188–216.
27. Olive K.A., Steigman G. and Walker T.P. Primordial nucleosynthesis: Theory and observations. *Phys. Rep.*, 2000, v. 333–334, 389–407.
28. Pagel B.E.J. Helium and Big Bang Nucleosynthesis. *Phys. Rep.*, 2000, v. 333–334, 433–447.
29. Steigman G. Primordial nucleosynthesis: Successes and challenges. *Int. J. Mod. Phys. E*, 2006, v. 15, 1–36.
30. Peimbert M. The primordial helium abundance. *Curr. Sci.*, 2008, v. 95, 1165–1176; arXiv: astro-ph/0811.2980.
31. Greenstein G.S., Truran J.W. and Cameron A.G.W. Helium deficiency in old halo B type stars. *Nature*, 1967, v. 213, 871.
32. Lindemann F.A. LXX. Note on the theory of magnetic storms. *Phil. Mag.*, 1919, v. 38(228), 669–684.
33. Eggert. J. Über den Dissoziationszustand der Fixsterngase. *Physikalische Zeitschrift*, 1919, v. 20, 570–574.
34. Kothari D.S. Meghnad Saha. 1893–1956. *Biogr. Mems. Fell. R. Soc.*, 1960, v. 5, 216–236.
35. Chatterjee S. Meghnad Saha — The scientist and the institution builder. *Ind. J. Hist. Sci.*, 1994, v. 29(1), 99–110; http://www.new.dli.ernet.in/rawdataupload/upload/insa/INSA_1/20005b68_99.pdf

36. Saha M.H. LIII. Ionization in the solar chromosphere. *Phil. Mag. Series 6*, 1920, v. 40, 472–488.
37. Saha M.H. On the physical theory of stellar spectra. *Proc. Roy. Soc. London A*, 1921, v. 99, 135–153.
38. Nernst W. Die theoretische un experimentellen Grundlagen des neuen Wärmesatzes. Verlag von Wilhelm Knapp, Halle, 1918, p.154.
39. Clayton D.D. Principles of stellar evolution and nucleosynthesis. Mcgraw-Hill, New York, N.Y., 1968.
40. Kippenhahn R. and Weigert A. Stellar structure and evolution. Springer-Verlag, Berlin, 1990.
41. Gingerich O. Obituary — Payne-Gaposchkin Cecilia. *Quart. J. Roy. Astron. Soc.*, 1982, v. 23, 450–451.
42. McCrea W. Obituary: Subramanyan Chandrasekhar. *The Observatory*, 1996, v. 116, 121–124.
43. Chandrasekhar S. Ralph Howard Fowler. *Astrophys. J.*, 1945, v.101(1), 1–5.
44. Taylor R.J. E.A. Milne(1896-1950) and the structure of stellar atmospheres and stellar interiors. *Quart. J. Roy. Astron. Soc.*, 1996, v. 37, 355–363.
45. Fowler R.H. and Milne E.A. The intensities of absorption lines in stellar spectra, and the temperature and pressures in the reversing layers of stars. *Mon. Not. Roy. Astron. Soc.*, 1923, v. 83, 403–424.
46. Fowler R.H. and Milne E.A. The maxima of absorption lines in stellar spectra (Second Paper). *Mon. Not. Roy. Astron. Soc.*, 1923, v. 83, 403–424.
47. Milne E.A. XVIII. Statistical equilibrium in relation to the photo-electric effect, and its application the determination of the absorption coefficients. *Phil. Mag. Ser. 6*, 1924, v. 47, 209–241.
48. Milne E.A. Ionization in stellar atmospheres. Part 1. Generalised Saha formulae, maximum intensities, and the determination of the coefficient of opacity. *Mon. Not. Roy. Astron. Soc.*, 1928, v. 89, 17–49.
49. Morgan W.W., Keenan P.C. and Kellman E. An atlas of stellar spectra. University of Chicago Press, Chicago, Il, 1943.
50. Gingerich O. Theory and observation of normal stellar atmospheres: Proceedings of the third Harvard-Smithsonian conference on stellar atmospheres. The M.I.T. Press, Cambridge, M.A., 1969.
51. Mitchell H.B. Henrietta Leavitt and the Cepheid variables. *The Physics Teacher*, 1976, v. 14(3), 162–167.
52. Bok P.F. Annie Jump Cannon 1863–1941. *Publ. Astron. Soc. Pac.*, 1941, v. 53, 168–170.
53. Merrill P. Annie Jump Cannon. *Mon. Not. Roy. Astron. Soc.*, 1942, v. 102(2), 74–76.
54. Gaposchkin C.P. Annie Jump Cannon. *Science*, 1941, v. 93(2419), 443–444.
55. Hoffleit D. Antonia C. Maury. *Sky & Telescope*, 1952, v. 11(5), 106.
56. Cannon A.J. Williamina Paton Fleming. *Science*, 1911, v. 33(861), 987–988.
57. Cannon A.J. Williamina Paton Fleming. *Astrophys. J.*, 1911, v. 34, 314–317.
58. Spradley J.L. The Industrious Mrs. Fleming. *Astronomy*, 1990, v. 18(7), 48–51.
59. Spradley J.L. Two centennials of star catalogs compiled by women. *Astron. Quart.*, 1990, v. 7(3), 177–184.
60. Lankford J. and Slavings R.L. Gender and Science: Women in American Astronomy, 1859–1940. *Phys. Today*, 1990, v. 43(3), 58–65.
61. Merton T.R. and Barratt S. Bakerian Lecture: On the spectrum of hydrogen. *Phil. Trans. Roy. Soc. London*, 1922, v. 222, 369–400.
62. Lang K.R. and Gingerich O. A source book in astronomy and astrophysics, 1900–1975. Harvard University Press, Cambridge, MA, 1979, Chap. 39, p.244.
63. McCrea W.H. The hydrogen chromosphere. *Mon. Not. Roy. Astron. Soc.*, 1929, v. 89, 483–497.
64. Przybilla N. and Butler K. The solar hydrogen spectrum in non-local thermodynamic equilibrium. *Astrophys. J.*, 2004, v. 610, L61–L24.
65. Robitaille P.-M. On the Presence of a Distinct Solar Surface: A Reply to Hervé Faye. *Prog. Phys.*, 2011, v. 3, 75–78.
66. Robitaille P.-M. Magnetic Fields and Directional Spectral Emissivity in Sunspots and Faculae: Complimentary Evidence of Metallic Behavior on the Surface of the Sun. *Prog. Phys.*, 2013, v. 1, 19–24.
67. Moore C.E. Henry Norris Russell. *The Observatory*, 1957, v. 77, 67–68.
68. de Jager C., Migeotte M., and Neven L. The profile of the Brackett alpha line in the solar spectrum. *Ann. Astrophys.*, 1956, v. 19, 9–18.
69. Russell H.N. and Stewart J.Q. Pressures at the Sun's surface. *Astrophys. J.*, 1924, v. 59, 197–209.
70. Rudnick P. A quantitative investigation of spectral line intensities in O- and B-type stars. *Astrophys. J.*, 1936, v. 83, 439–475.
71. Underhill A.B. and Petrie W. The stark effect of helium in some B type stars. *J. Roy. Astron. Soc. Canada*, 1944, v. 38, 385–394.
72. Underhill A.B. The relative abundance of hydrogen to helium in stars. *J. Roy. Astron. Soc. Canada*, 1953, v. 47, 153–159.
73. Underhill A.B. On the strength of the helium line in O-type stars. Les Processus Nucléaires dans les Astres. Communications présentées au cinquième *Colloque International d'Astrophysique* tenu à Liège les 10-12 Septembre, 1953, 374–386.
74. Milne E.A. Selective radiation-pressure and the structure of a stellar atmosphere. *Mon. Not. Roy. Astron. Soc.*, 1927, v. 87, 697–708.
75. Milne E.A. The effect of collisions on monochromatic radiative equilibrium. *Mon. Not. Roy. Astron. Soc.*, 1928, v. 88, 493–502.
76. Milne E.A. Bakerian Lecture: The structure and opacity of a stellar atmosphere. *Phil. Trans. Roy. Soc. London*, 1929, v. 228, 421–461.
77. Milne E.A. Thermodynamics of the stars. Handbuch der Astrophysik, 1930, v.3, Part 1, 65–255 (also in Menzel D.H. Selected Papers on the Transfer of Radiation: 6 papers by Arthur Schuster, K. Schwarzschild, A.S. Eddington, S. Rosseland, E.A. Milne. Dover Publications, Inc., New York, 1966, 77–269).
78. Robitaille P.M. Stellar opacity: The Achilles' heel of the gaseous Sun. *Progr. Phys.*, 2011, v. 3, 93–99.
79. Gerasimovic B.P. On the correction to Saha's formula for small deviations from thermodynamic equilibrium. *Proc. Nat. Acad. Sci. USA*, 1927, v. 13(4), 180–185.
80. Noyes A.A. and Wilson H.A. The thermal ionization of gaseous elements at high temperatures: a confirmation of Saha theory. *Astrophys. J.*, 1923, v. 57, 20–32.
81. van der Mullen J.A.M., Benoy D.A., Fey F.H.A.G. and van der Sijde B. Saha equation for two-temperature plasmas: Theories, experimental evidence, and interpretation. *Phys. Rev. B*, 1994, v. 50(5), 3925–3934.
82. Sweeney M.A. Thermodynamic inconsistency of the modified Saha equation at high pressures. *Astrophys. J.*, 1978, v. 220, 335–338.
83. Pottasch S.R. and Thomas R.N. Departures from the Saha equation under varying conditions of Lyman continuous opacity. *Astrophys. J.*, 1959, v. 130, 941–953.
84. Krawec R. Steady-state composition of low-density nonequilibrium hydrogen plasma. NASA Technical Note D-3457, Washington, D.C., October 1966.
85. Kurochka L.N. Saha's equation under deviation from thermodynamic equilibrium. *Bull. Astron. Inst. Czechoslovakia*, 1973, v. 24(4), 210–212.
86. Avrett E.H. Solution of non-LTE transfer problems. *J. Quant. Spectrosc. Radiat. Trans.*, 1971, v. 11(6), 511–529.

87. Canal G.P., Luna H., Galvão R.M.O. and Castell R. An approach to a non-LTE Saha equation based on Druyvesteyn energy distribution function: A comparison between electron temperature obtained from OES and the Langmuir probe analysis. *J. Phys. D*: *Appl. Phys.*, 2009, v. 42, 135202 (6 pages).
88. Auer L.H. The stellar atmosphere problem. *J. Quant. Spectrosc. Radiat. Transfer*, 1971, v. 11, 573–587.
89. Fridman A.A. and Kennedy L.A. Plasma physics and engineering. Taylor & Francis, New York, N.Y., 2004.
90. Mihalas D. Stellar atmospheres (2nd Edition). W.H. Freeman and Company, San Francisco, CA, 1978
91. Hubeny I. Theory and modeling of stellar atmospheres. *AIP Conf. Proc.*, 2010, v. 1268, 73–115.
92. Werner K. and Dreizler S. The classical stellar atmosphere problem. *J. Comp. Appl. Math.*, 1999, v. 109(1–2), 65-93.
93. Lecar M. Departures from local thermodynamic equilibrium in an A0 stellar atmosphere. NASA Technical Note, D 2110, 1964.
94. Przybilla N., Nieva M.-F. and Butler K. Testing common classical LTE and NLTE model atmosphere and line-formation codes for quantitative spectroscopy of early-type stars. *J. Phys. Conf. Ser.*, 2011, v. 328, 012015 (12 pages).
95. Nieva M.F. and Przybilla N. Hydrogen and helium line formation in OB dwarfs and giants: A hybrid non-LTE approach. *Astron. Astrophys.*, 2007, v. 467, 295–309.
96. Eddington A.S. Further notes on the radiative equilibrium of the stars. *Mon. Not. Roy. Astron. Soc.*, 1917, v. 77, 596–612.
97. Eddington A.S. On the radiative equilibrium of the stars. *Mon. Not. Roy. Astron. Soc.*, 1916, v. 77, 16–35.
98. Eddington A.S. On the absorption of radiation inside a star. *Mon. Not. Roy. Astron. Soc.*, 1922, v. 83, 32–46.
99. Kramers H.A. XCIII. On the theory of x-ray absorption and of the continuous x-ray spectrum. *Phil. Mag.*, 1923, 46(275), 836–871.
100. Strömgren B. The opacity of stellar matter and the hydrogen content of the stars. *Zeitschrift für Astrophysik*, 1932, v. 4, 118–152.
101. Strömgren B. On the interpretation of the Hertzaprung-Russell-Diagram. *Zeitschrift für Astrophysik*, 1933, v. 7, 222–248.
102. Wildt R. Electron affnity in astrophysics. *Astrophys. J.*, 1939, v. 89, 295–301.
103. Cowling T.G. The stability of gaseous stars. *Mon. Not. Roy. Astron. Soc.*, 1935, v. 96, 42–60.
104. Schwarzschild M. On the helium content of the Sun. *Astrophys. J.*, 1946, v. 104, 203–207.
105. Iinuma Y. On the determination of chemical composition of stars. *Publ. Astron. Soc. Japan*, 1949, v. 1(1–2), 18–23.
106. Schwarzschild M., Howard R. and Härm R. Inhomogeneous stellar models. V. A solar model with convective envelope and inhomogeneous interior. *Astrophys. J.*, 1957, v. 125, 233–241.
107. Weymann R. Inhomogeneous stellar models. VI. An improved solar model with the carbon cycle included. *Astrophys. J.*, 1957, v. 126, 208–212.
108. Osterbrock D.E. and Rogerson J.B. The helium and heavy-element content of gaseous nebulae and the Sun. *Publ. Astron. Soc. Pac.*, 1961, v. 73, 129–134.
109. Bahcall J.N., Fowler W.A., Iben I. and Sears R.L. Solar neutrino flux. *Astrophys. J.*, 1963, v. 137, 344–346.
110. Carson T.R. Stellar opacity. *Ann. Rev. Astron. Astrophys.*, 1976, v. 14, 95–117.
111. Rozsnyai B.F. Solar opacities. *J. Quant. Spec. Rad. Trans.*, 2001, v. 71, 655–663.
112. The Opacity Project Team. The Opacity Project. Institute of Physics Publishing, Bristol, UK, 1995, v.1.
113. The Opacity Project Team. The Opacity Project. Institute of Physics Publishing, Bristol, UK, 1996, v.2.
114. Rogers F.J. and Iglesias C.A. Stellar opacity. Lawrence Livermore National Laboratory, Preprint UCLRL-JC-137066.
115. Iglesias C.A. and Rogers F.J. Opacities for the solar radiative interior. *Astrophys. J.*, 1991, v. 371, 408–417.
116. Iglesias C.A. and Rogers F.J. Radiative opacities for carbon- and oxygen-rich mixtures. *Astrophys. J.*, 1993, v. 412, 752–760.
117. Rogers F.J. and Iglesias C.A. Rosseland mean opacities for variable compositions. *Astrophys. J.*, 1992, v. 401, 361–366.
118. Iglesias C.A. and Rogers F.J. Updated OPAL opacities. *Astrophys. J.*, 1996, v. 464, 943–953.
119. Aller L.H. and Chapman S. Diffusion in the Sun. *Astrophys. J.*, 1960, v. 132, 461–472.
120. Noerdlinger P.D. Diffusion of helium in the Sun. *Astron. Astrophys.*, 1977, v. 57, 407–415.
121. Aldo M. Serenelli A.M. and Basu S. Determining the initial helium abundance of the Sun. *Astrophys. J.*, 2010, v. 719(1), 865–872.
122. Alpher R.A., Bethe H., and Gamow G. The origin of chemical elements. *Phys. Rev.*, 1948, v. 73(7), 803–804.
123. Hoyle F. and Taylor R.J. The Mystery of the Cosmic Helium Abundance. *Nature*, 1964, v. 203(4950), 1108–1110.
124. Wagoner R.V., Fowler W.A. and Hoyle F. On the synthesis of elements at very high temperatures. *Astrophys. J.*, 1967, v. 148, 3–49.
125. Burbidge G. Cosmic helium. *Comm. Astrophys. Space Phys.*, 1969, v. 1, 101–106.
126. Penzias A. A. and Wilson R.W. A measurement of excess antenna temperature at 4080 Mc/s. *Astrophys. J.*, 1965, v. 1, 419–421.
127. Hubble E. A relation between distance and radial velocity among extragalactic nebulae. *PNAS*, 1929, v. 15(3), 168–173.
128. Rabounski D. On the Exact Solution Explaining the Accelerate Expanding Universe According to General Relativity. *Progr. Phys.*, 2012, v. 2, L1–L6.
129. Lemaître G. Un Univers homogène de masse constante et de rayon croissant rendant compte de la vitesse radiale des nébuleuses extragalactiques. *Annales de la Societé Scientifique de Bruxelles*, ser.A, 1927, v. 47, 49–59.
130. Dicke R.H., Peebles P.J.E., Roll P.G., and Wilkinson D.T. Cosmic black-body radiation. *Astrophys. J.*, 1965, v. 1, 414–419.
131. Robitaille J.C. and Robitaille P.M. Liquid metallic hydrogen III. Intercalation and lattice exclusion versus gravitational settling, and their consequences relative to internal structure, surface activity, and solar winds in the Sun. *Progr. Phys.*, 2013, v. 2, in press.
132. Komatsu E., Smith K.M, Dunkley J., Bennett C.L., Gold B., Hinshaw G., Jarosik N., Larson D., Nolta M.R., Page L., Spergel D.N., Halpern M., Hill R.S., Kogut A., Limon M., Meyer S.S., Odegard N., Tucker G.S., Weiland J.L., Wollack W. and Wright E.L. Seven-year wilkinson microwave anisotropy probe (WMAP*) observations: Cosmological interpretation. *Astrophys. J. Suppl. Ser.*, 2011, v. 192, 18 (47 pp).
133. Hinshaw G., Larson D., Komatsu E., Spergel D.N., Bennett C.L., Dunkley J., Nolta M.R., Halpern M., Hill R.S., Odegard N., Page L., Smith K.M., Weiland J.L., Gold B., Jarosik N., Kogut A., Limon M., Meyer S.S., Tucker G.S., Wollack E. and Wright E.L. Nine-year wilkinson microwave anisotropy probe (WMAP) observations: Cosmological parameter results. arXiv: astro-ph.CO/1212.5226.
134. Pagel B.E.J, Simonson E.A., Terlevich R.J. and Edmunds M.J. The primordial helium abundance from observations of extragalactic H II regions. *Mon. Not. Roy. Astron. Soc.*, 1992, v. 255, 325–345.
135. Peimbert M., Peimbert A. and Ruiz M.T. The chemical composition of the Small Magellanic Cloud H II region NGC 346 and the primordial helium abundance. *Astrophys. J.*, 2000, v. 541, 688–700.

136. Luridiana V., Peimbert A., Peimbert M. and Cerviño M. The effect of collisional enhancement of Balmer lines on the determination of the primordial helium abundance. *Astrophys. J.*, 2003, v. 592, 846–865.

137. Izotov Y.I. and Thuan T.X. Systematic effects and a new determination of the primordial abundance of 4He and dY/dZ from observations of blue compact galaxies. *Astrophys. J.*, 2004, v. 602, 200–230.

138. Robitaille P.M. WMAP: A radiological analysis. *Progr. Phys.*, 2006, v. 1, 3–18.

139. Planck Satellite website: http://www.rssd.esa.int/index.php?project=planck

140. Olive K.A. and Steigman G. On the abundance of primordial helium. *Astrophys. J. Suppl. Ser.*, 1995, v. 97, 49–58.

141. Spitzer L. and Savedoff M.P. The temperature of interstellar matter III. *Astrophys. J.*, 1950, v. 111, 593–608.

142. Dicke R.H., Peebles P.J.E., Roll P.G., and Wilkinson D.T. Cosmic black-body radiation. *Astrophys. J.*, 1965, v. 1, 414–419.

143. Formisano V. and Moreno G. Helium and heavy ions in the solar winds. *Revista del Nuovo Cimento*, 1971, v. 1(3), 365–422.

144. Robbins D.E., Hundhausen A.J. and Bame S.J. Helium in the solar wind. *J. Geophys. Res.*, 1970, v. 75(7), 1178–1187.

145. Ogilvie K.W. and Hirshberg J. The solar cycle variation of the solar wind helium abundance. *J. Geophys. Res.*, 1974, v. 79(31), 4595–4602.

146. Bame S.J., Asbridge J.R., Feldman W.C. and Gosling J.T. Evidence for a structure-free state at high solar wind speeds. *J. Geophys. Res.*, 1977, v. 82(10), 1487–1492.

147. Borrini G., Gosling J.T., Bame S.J., and Feldman W.C. Helium abundance enhancements in the solar wind. *J. Geophys. Res.*, 1982, v. 87(A9), 7370–7378.

148. Bochsler P. Structure of the solar wind and compositional variations. *Space Sci. Rev.*, 1998, v. 85, 291–302.

149. Aellig M.R., Lazarus A.J. and Steinberg J.T. The solar wind helium abundance: Variations with wind speed and solar cycle. *Geophys. Res. Let.*, 2001, v. 28(14), 2767–2770.

150. Wurz P. Solar Wind Composition. In: *The Dynamic Sun: Challenges for Theory and Observations*, ESA SP-600, 2005, v. 5.2, 1–9.

151. Kasper J.C., Stevens M.L., Lazarus A.J. and Ogilvie K.W. Solar wind and helium abundance as a function of speed and heliographic latitude: Variation through a solar cycle. *Astrophys. J.*, 2007, v. 660, 901–910.

152. Wang Y.M. Relating the solar wind helium abundance to the coronal magnetic field. *Astrophys. J.*, 2008, v. 683, 499–509.

153. Geiss J. and Bochsler P. Long time variations in solar wind properties: possible causes versus observations. In: *The Sun in Time* (Sonett C.P., Giampapa M.S., Matthews M.S., eds.), The University of Arizona Press, Tucson, AZ, 1991, p.98-117.

154. von Steiger R., Vial J.C., Bochsler P., Chaussidon M., Cohen C.M.S., Fleck B., Heber V.S., Holweger H., Issautier K., Lazarus A.J., Ogilvie K.W., Paquette J.A., Reisenfeld D.B, Teriaca L., Wilhelm K., Yusainee S., Laming J.M., and Wiens R.C. Measuring solar abundances. *AIP Conf. Proc.*, 2001, v. 598, 13–22.

155. Grevesse N. and Sauval A.J. Standard solar composition. *Space Science Reviews*, 1998, v. 85, 161–174.

156. Asplund M., Grevesse N., Sauval A.J. and Scott P. The chemical composition of the Sun. *Ann. Rev. Astron. Astrophys.*, 2009, v. 47, 481–522.

157. Robitaille P.M. Liquid Metallic Hydrogen: A Building Block for the Liquid Sun. *Progr. Phys.*, 2011, v. 3, 60–74.

New Constraints on Quantum Theories

Eliahu Comay

Charactell Ltd., PO Box 39019, Tel-Aviv, 61390, Israel. E-mail: elicomay@post.tau.ac.il

Hierarchical relationships between physical theories are discussed. It is explained how a lower rank theory imposes constraints on an acceptable structure of its higher rank theory. This principle is applied to the case of quantum mechanics and quantum field theory of massive particles. It is proved that the Dirac equation is consistent with these constraints whereas the Klein-Gordon equation, as well as all other second order quantum equations are inconsistent with the Schrödinger equation. This series of arguments undermines the theoretical structure of the Standard Model.

1 Introduction

The equations of motion are regarded as the basis of a physical theory. A mathematical analysis of these equations yields the complete form of a given theory and of its details. The validity of a mathematically correct physical theory should be consistent with two kinds of tests. Thus, it must agree with relevant experimental data and it must also be consistent with well established physical principles. (Evidently, the latter represent many experimental data in a concise form.) The following simple example illustrates the latter point. A new theory is unacceptable if its final results are inconsistent with the law of energy conservation. This point shows the significance of physical constraints that restrict the number of acceptable physical theories and guide theoretical and experimental efforts to take promising directions.

The definition of a domain of validity is an important element of a theory. For example, mechanics is the science used for predicting the motion of bodies. It is very successful in the case of the motion of planets moving around the sun. On the other hand, it cannot predict the motion of an eagle flying in the sky. This example does not mean that mechanics is incorrect. It means that mechanics is a very satisfactory science for a set of experiments. For example, Newtonian mechanics is acceptable for cases where the following conditions hold: the velocity is much smaller than the speed of light, the classical limit of quantum mechanics holds, and the force can be calculated in terms of position, time and velocity. The set of experiments where a given theory is successful is called the theory's domain of validity. This issue is used in the rest of this work.

The definition of the domain of validity illustrates an important aspect of the correctness of a physical theory. Indeed, this notion should be regarded in a relative sense. Thus, many measurements are given together with experimental error. For this reason, even if we know that a given theory is not perfect, it still can be regarded as a correct theory for cases where the theory's errors are smaller than the experimental errors.

In this work units where $\hbar = c = 1$ are used. In this system of units one kind of dimension applies and here it is the length $[L]$. Thus, the dimension of every physical quantity takes an appropriate power of $[L]$. For example, mass, energy and momentum take the dimension $[L^{-1}]$. The metric is diagonal and its entries are $(1, -1, -1, -1)$. Greek indices run from 0 to 3. The subscript symbol $_{,\mu}$ denotes the partial differentiation with respect to x^{μ}.

2 The dimensions of quantum fields

Consider the two sets of experiments S_A and S_B defining the domains of validity of the physical theories A and B, respectively.

Fig. 1 illustrates the hierarchical relationships between theories A and B. Here the sets S_A and S_B consist of all experiments that are described correctly by theory A and B, respectively. The set S_A is a subset of S_B. This relationship means that all experiments that are described successfully by theory A are also described successfully by theory B, but not vice versa. For this reason it can be stated that theory B has a more profound meaning because it is also valid for cases where theory A is useless. However, this fact does not mean that theory A is wrong, simply because this theory can be used successfully for all cases that belong to its domain of validity S_A.

This kind of relationships between theories has been recognized a long time ago. For example, A. Einstein mentions special relativity and general relativity and explains why special relativity should not be regarded as a wrong theory. The reason is that special relativity holds in cases where a flat space-time can be regarded as a good description of the

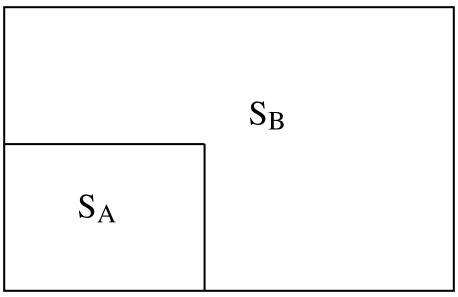

Fig. 1: Domains of validity of two theories (see text).

physical conditions. Similarly, considering electrostatics and Maxwellian electrodynamics, he explains why electrostatics is a good theory for cases where the charge carriers can be regarded as motionless objects (see [1], pp. 85, 86).

The issue of hierarchical relationships between theories is also discussed in Rohrlich's book (see [2], pp. 1–6). Here one can find explanation showing the hierarchical relationships between several pairs of theories. This discussion provides the reader with a broader overview of the structure of existing physical theories and of their hierarchical relationships.

As pointed out above, a physical theory that takes a higher hierarchical position has a more profound meaning. The rest of this work relies on another result obtained from these relationships. Thus, *a well established physical theory imposes constraints on appropriate limits of a higher rank theory.* For example, this requirement is satisfied by relativistic mechanics, whose low velocity limit agrees with Newtonian mechanics (see [3], pp. 26–30). Similarly, the classical limit of quantum mechanics agrees with classical physics (see [4], pp. 19–21 and [5], pp. 133–141). Below, this principle is called *constraints imposed by a lower rank theory.* It is shown in this work that this principle provides powerful constraints on the acceptability of physical theories.

3 Hierarchical Relationships Between Quantum Theories

Let us discuss the hierarchical relationships between three quantum theories of massive particles: non-relativistic quantum mechanics (QM), relativistic quantum mechanics (RQM) and quantum field theory (QFT) (see fig. 2). Thus, QM takes the lowest hierarchical rank because it is valid for cases where the absolute value of the momentum's expectation value is much smaller than the particle's self-mass. RQM is valid for cases where the number of particles can be regarded as a constant of the motion. QFT is a more general theory and RQM is its appropriate limit. The inherent relationships between these theories are well documented in the literature. Thus, S. Weinberg makes the following statement. "First, some good news: quantum field theory is based on the same quantum mechanics that was invented by Schrödinger, Heisenberg, Pauli, Born, and others in 1925-1926, and has been used ever since in atomic, molecular, nuclear and condense matter physics" (see [6], p. 49).

The Schrödinger equation takes the following form

$$i\frac{\partial\psi}{\partial t} = -\frac{1}{2m}\Delta\psi + U\psi. \tag{1}$$

An analysis of this equation yields an expression for a conserved current whose density is (see e.g. [4], pp. 53–55)

$$\rho = \psi^*\psi. \tag{2}$$

Relation (2) proves that the dimension of the Schrödinger function is

$$[\psi] = [L^{-3/2}]. \tag{3}$$

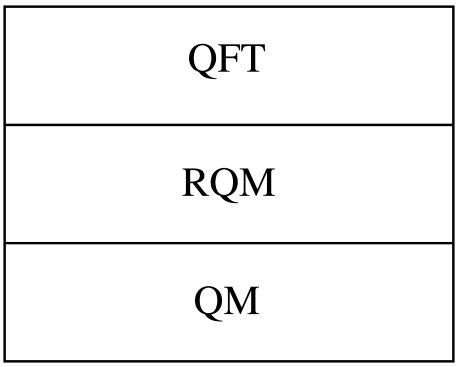

Fig. 2: Hierarchical relationships between three quantum theories (see text).

Here the expression for density depends only on the wave function and contains no derivatives. The form of the density (2) is an important element of the theory because it enables a construction of a Hilbert space of the time-independent functions which belong to the Heisenberg picture.

Let us examine the structure of QFT. The vital role of the Lagrangian density in QFT can be briefly described as follows. The phase is an indispensable element of quantum theories. Being an argument of an exponent which can be expanded in a power series, the phase must be a dimensionless Lorentz scalar. Thus, the phase is defined as a Lorentz scalar action (divided by $\hbar$). The following expression shows how the action is obtained from a given Lagrangian density $\mathcal{L}$

$$S = \int \mathcal{L}\, d^4x. \tag{4}$$

This expression proves that a dimensionless Lorentz scalar action is obtained from a Lagrangian density that is a Lorentz scalar whose dimension is $[L^{-4}]$.

This property of the Lagrangian density is used in an examination of two kinds of QFT theories. Let us begin with the first order Lagrangian density of a free Dirac field ψ_D (see [7], p. 54)

$$\mathcal{L}_D = \bar{\psi}_D[\gamma^\mu i\partial_\mu - m]\psi_D. \tag{5}$$

Now, the dimension $[L^{-4}]$ of the Lagrangian density and the dimension $[L^{-1}]$ of the operators ∂_μ and m prove that the dimension of the Dirac field ψ_D is $[L^{-3/2}]$. This value agrees with that of the Schrödinger function (3). It means that the Dirac field theory satisfies the dimension constraints imposed by the lower rank theory of QM.

A different result is obtained from the second order complex Klein-Gordon (KG) equation. The Lagrangian density of this equation is (see [7], p. 38)

$$\mathcal{L}_{KG} = g^{\mu\nu}\phi^*_{,\mu}\phi_{,\nu} - m^2\phi^*\phi. \tag{6}$$

Here the dimension of the operators is $[L^{-2}]$. Using the dimension $[L^{-4}]$ of the Lagrangian density, one infers that the

dimension of the KG function ϕ is $[L^{-1}]$. On the other hand, it is shown in (3) that the dimension of the Schrödinger wave function is $[L^{-3/2}]$. This outcome means that the complex KG function ϕ violates a constraint imposed by a lower rank theory.

It turns out that this inconsistency holds for other quantum equations where the dimension of their field function is $[L^{-1}]$. Thus, a dimension $[L^{-1}]$ is a property of the following field: the Yukawa particle (see [8], p. 211), the electroweak $W^{\pm}, Z$ bosons (see [9], p. 307) and the Higgs boson (see [10], p. 715). For this reason, quantum theories of all these particles are inconsistent with the dimensional constraint imposed by the Schrödinger equation.

One can also see immediately that the Yukawa and the Z fields introduce to the Lagrangian density an interaction term with a fermion ψ which takes the form

$$\mathcal{L}_{Int} = g\bar{\psi}\phi\psi. \tag{7}$$

This kind of interaction means that the field ϕ of each of these particles is a *real* field (in a mathematical sense). This conclusion stems from the facts that the action and the integration factor d^4x are real. These properties mean that all terms of a Lagrangian density must be real. Now, since g and the product $\bar{\psi}\psi$ are real, one finds that ϕ is real. Evidently, a theory of a real field is inconsistent with another constraint of QM. Indeed, QM uses a *complex* wave function and for this reason the non-relativistic limits of the *real* field of Yukawa and of Z particles also violate a second kind of constraint.

4 Concluding Remarks

It is explained in this work how hierarchical relationships between physical theories can be used for deriving necessary conditions that an acceptable higher rank theory must satisfy. This issue is applied to QFT theories and the non-relativistic limit of their field function is compared with properties of non-relativistic quantum mechanics. It is explained how such a comparison provides a powerful criterion for the acceptability of physical theories. The discussion examines the dimension of quantum functions of several specific theories and compares the dimension of QFT theories with that of the lower rank non-relativistic Schrödinger theory. It turns out that the Dirac field satisfies this criterion whereas the Klein-Gordon and the Yukawa theories as well as those of the $W^{\pm}, Z$ and the Higgs boson fail to satisfy this criterion.

An important evaluation of a theoretical idea is a comparison of its outcome with experimental results. Referring to this issue, one should note that a field function $\psi(x^{\mu})$ which is used in QM, RQM and QFT depends on a *single* set of four space-time coordinates x^{μ}. For this reason, $\psi(x^{\mu})$ describes an elementary point-like particle. The following example illustrates this matter. A pion consists of a quark-antiquark pair of the u, d flavor and each quark is described by a function that depends on its own 4-coordinates x^{μ}. Hence, a pion cannot be described by a function $\psi(x^{\mu})$, simply because this function has a smaller number of independent coordinates. It turns out that experimental data of all spin-1/2 Dirac particles, namely, leptons and quarks, are consistent with their pointlike attribute. On the other hand the pion, which was the original KG candidate is not pointlike and the $\pi^{\pm}$ mesons have a charge radius which is not much smaller than that of the proton [11]. There is still no experimental data concerning pointlike properties of the $W^{\pm}, Z$ and the Higgs boson.

As is well known, the $W^{\pm}, Z$ and the Higgs bosons are cornerstones of the Standard Model. It means that the series of arguments presented in this work undermines the theoretical structure of the Standard Model. Evidently, a physical theory that has an inconsistent structure is unacceptable. Hence, people who still adhere to the Standard Model must show why the arguments presented above are incorrect. It is also interesting to note that the results of this work are consistent with Dirac's lifelong objection to the second order KG equation of a spin-0 boson (see [12], pp. 3, 4).

Submitted on December 21, 2012 / Accepted on January 4, 2013

References

1. Einstein A. Albert Einstein in His Own Words. Portland House, New York, 2000.
2. Rohrlich F. Classical Charged Particle. World Scientific, New Jersey, 2007.
3. Landau L.D. and Lifshitz E.M. The Classical Theory of Fields. Elsevier, Amsterdam, 2005.
4. Landau L.D. and Lifshitz E.M. Quantum Mechanics. Pergamon, London, 1959.
5. Schiff L.I. Quantum Mechanics. McGraw-Hill, New York, 1955.
6. Weinberg S. The Quantum Theory of Fields, Vol. I, Cambridge University Press, Cambridge, 1995.
7. Bjorken J.D. and Drell S.D. Relativistic Quantum Fields. McGraw-Hill, New York, 1965.
8. Bjorken J.D. and Drell S.D. Relativistic Quantum Mechanics. McGraw-Hill, New York, 1964.
9. Weinberg S. The Quantum Theory of Fields, Vol. II, Cambridge University Press, Cambridge, 1996.
10. Peskin M.E. and Schroeder D.V. An Introduction to Quantum Field Theory. Addison-Wesley, Reading (Mass), 1995.
11. Beringer J. et al. REVIEW OF PARTICLE PHYSICS Particle Data Group. *Physical Review D*, 2012, v. 86, 010001.
12. Dirac P.A.M. In: Mathematical Foundations of Quantum Theory, Editor A.R. Marlow, Academic, New York, 1978.

The Nuclear Shape Phase Transitions Studied within the Geometric Collective Model

Khalaf A.M.* and Ismail A.M.†

*Physics Department, Faculty of Science, Al-Azhar University, Cairo, Egypt. E-mail: ali_khalaf43@hotmail.com
†Hot Laboratories Center, Atomic Energy Authority, Egypt, P.No. 13759, Cairo, Egypt. E-mail: ahmedismailph@yahoo.com

In the framework of the Geometric Collective Model (GCM), quantum phase transition between spherical and deformed shapes of doubly even nuclei are investigated. The validity of the model is examined for the case of lanthanide chains Nd/Sm and actinide chains Th/U. The parameters of the model were obtained by performing a computer simulated search program in order to obtain minimum root mean square deviations between the calculated and the experimental excitation energies. Calculated potential energy surfaces (PES's) describing all deformation effects of each nucleus are extracted. Our systematic studies on lanthanide and actinide chains have revealed a shape transition from spherical vibrator to axially deformed rotor when moving from the lighter to the heavier isotopes.

1 Introduction

The nuclear shape transitions were studied within the nuclear interacting boson model (IBM) [1–3]. The IBM-1 describes a system of a fixed number N of spin zero and two bosons (s and d bosons) subject to one- and two-body interactions. The IBM-1 reveals a transparent algebraic structure with U(6) as the dynamical group. Varying six free parameters of the model, one can reach three standard dynamical symmetries U(5), SU(3) and O(6) and two additional ones SU(3)* and O(6)* [2]. It turns out that these dynamical symmetries provide an appropriate framework for the description of low-energy collective motions of real nuclei with certain shape symmetries: The U(5) limit corresponds to spherical nuclei, the SU(3) and SU(3)* limits to axially symmetric nuclei with quadruple deformation (prolate and oblate shapes) and the O(6) and O(6)* limits to quadruply deformed nuclei that are unstable against the axial symmetry breaking. This is represented in the so called Casten triangle [2,4] with vertices corresponding to the standard dynamical symmetries and the other points to various transitional cases. Phase transitions between these shapes were studied, and it is known that the phase transition from U(5) to O(6) is second order, while any other transition within the Casten triangle from a spherical to a deformed shape is first order [5–15].

Alternative descriptions of nuclei at the critical point of phase transitions from spherical vibrator to deformed γ soft E(5) [16], and from spherical vibrator to deformed axially symmetric rotor X(5) [17], were proposed. These analytic solutions are obtained by introducing a square well potential in the Bohr Hamiltonian and yield parameter free predictions for both energies and electromagnetic transition probabilities. Empirical examples were suggested for both the proposed symmetries [18]. It was found [19, 20] that the X(5) predictions cannot be exactly reproduced by any point in the two parameter space of the IBM, whereas best agreement is obtained for parameters corresponding to a point close to, but outside, the shape phase transition region of the IBM. Since the IBM was formulated from the beginning in terms of creation and annihilation boson operators, its geometric interpretation in terms of shape variables is usually done by introducing a boson condensate with two shape parameters β and γ through the intrinsic state formalism (coherent state) [21]. The parameter β is related to the axial deformation of the system, while γ measures the deviation from axial symmetry. The equilibrium shape of the system is obtained by minimizing the intrinsic state. It is well know that the dynamical symmetry associated with U(5) corresponds to a spherical shape $\beta = 0$, the dynamical symmetry SU(3) is associated with an axially deformed shape $\beta \neq 0$ and $\gamma = 0, \pi/3$ and the dynamical symmetry O(6) is related to a γ-unstable deformed shape $\beta \neq 0$ and γ-independent.

A very flexible and powerful approach to describe nuclear collective excitations which is an extension of the Bohr-Mottelson vibrational Hamiltonian [22] is the GCM essentially based on the quadruple degrees of freedom [23,24]. The problem of nuclear collective motion is formulated by Bohr and Mottelson from the beginning in terms of the intrinsic parameters β, γ and the three Euler angels ω_i that characterize the orientation of a deformed nucleus.

The GCM is a macroscopic nuclear structure model in the sense that it considers the nucleus as a charged liquid drop with a definite surface, rather than a many-body system of constituent particles.

Neodymium isotopes are the members of the chain of nuclei which represent an ideal case for studying the influence of the shape transition from spherical to deformed nuclei. Therefore, in the chart of nuclei there is a very important lanthanide Nd/Sm transition region which exhibit a rapid structural change from spherical to well deformed when moving from the lighter to the heavier isotopes. Although this tran-

sitional region has been studied extensively in the framework of the IBM, the discussion of phase transitions has not always been treated in a proper way.

In the present paper, we have analyzed systematically the transitional region and phase transition in lanthanide and actinide chains of isotopes in the framework of GCM. For each isotope chain a fitting procedure is performed to get the model parameters. We have generated the PES to classify phase transitions and to decide if a nucleus is close to criticality. In these chains, nuclei evolve from spherical to deformed shapes.

2 The GCM Hamiltonian and the PES's

The Hamiltonian of the GCM [23] represents a concrete realization of the general Bohr Hamiltonian [22] describing the quadruple oscillations of the nuclear surface. The collective Hamiltonian restricted to quadruple deformations can be written in the notation of Rajah for tensor products of irreducible tensor operators. The α's are the well known collective coordinates, which are defined by the usual expansion of the nuclear radius in terms of spherical harmonics. The $\hat{\pi}$ is the covariant tensor of the canonically conjugate momenta. We start by writing the GCM Hamiltonian as:

$$\hat{H} = \hat{T} + \hat{V}. \tag{1}$$

The kinetic energy $\hat{T}$ up to second order is given by [2].

$$\hat{T} = \frac{1}{B_2}[\pi \times \pi]^0 + \frac{P_3}{3}\left[[\pi \times \alpha]^{(2)} \times \hat{\pi}\right]^{(0)} \tag{2}$$

where B_2 is the common mass parameter and P_3 is an enharmonic kinetic term which for simplicity, we set to zero here. A transformation to the intrinsic body fixed system leads to a formal separation of the rotational and vibrational variables expressed by the Euler angles and the shape parameters β and γ respectively. The potential energy V is given by

$$\begin{aligned} V &= C_2[\alpha \times \alpha]^{(2)} + C_3\left[[\alpha \times \alpha]^{(2)} \times \alpha\right]^{(0)} + \\ &+C_4[\alpha \times \alpha]^{(0)}[\alpha \times \alpha]^{(0)} + \\ &+C_5[\alpha \times \alpha]^{(0)}\left[[\alpha \times \alpha]^{(2)} \times \alpha\right]^{(0)} + \\ &+C_6\left[[\alpha \times \alpha]^{(2)} \times \alpha\right]^{(0)}\left[[\alpha \times \alpha]^{(2)} \times \alpha\right]^{(0)} + \\ &+D_6[\alpha \times \alpha]^{(0)}[\alpha \times \alpha]^{(0)}[\alpha \times \alpha]^{(0)}. \end{aligned} \tag{3}$$

The six stiffness parameters C_2, C_3, C_4, C_5, C_6 and D_6 occurring in the collective potential energy are constants for each nucleus. They are treated as adjustable parameters which have to be determined from the best fit to the experimental data, level energies, B(E2) transition strengths and quadruple moments. They depend however on the proton and neutron numbers due to shell structure. The potential energy, expressed in terms of the intrinsic variables β and γ, is

$$\begin{aligned} V(\beta,\gamma) &= C_2\tfrac{1}{\sqrt{5}}\beta^2 - C_3 rub\sqrt{\tfrac{2}{35}}\beta^3\cos(3\gamma) + \\ &+ C_4\tfrac{1}{5}\beta^4 - C_5\sqrt{\tfrac{2}{175}}\beta^3\cos(3\gamma) + \\ &+ C_6\tfrac{2}{35}\beta^6\cos^2(3\gamma) + D_6\tfrac{1}{5\sqrt{5}}\beta^6 \\ &= V_s(\beta) + V_{Po}(\beta,\gamma) + V_{na}(\beta,\gamma). \end{aligned} \tag{4}$$

Roughly speaking the C_2, C_4 and D_6 terms describe the γ-independent features of the PES. They form the contribution $V_s(\beta)$. The C_3 and C_5 terms are responsible for the prolate-oblate energy differences in the PES and are represented by $V_{po}(\beta,\gamma)$. The C_6 term is symmetric about the $\gamma = \pi/6$ axis and therefore can be used for the generation of non axial shape $V_{na}(\beta,\gamma)$. The selection of the eight parameters of the GCM Hamiltonian is impractical and difficult, because the available observation data are usually not sufficient to establish the qualitative nature of the GCM potential. It is therefore, often desirable to use a more tractable form of the model. In practice simplification for the GCM is to use a maximum of three parameters to describe all limits of nuclear structure: vibrator, rotor and γ-soft nuclei and transition regions in between. Then the potential energy up to the fourth power of β is simplified to be:

$$V(\beta,\gamma) = C_2\frac{1}{\sqrt{5}}\beta^2 - C_3\sqrt{\frac{2}{35}}\beta^3\cos(3\gamma) + C_4\frac{1}{5}\beta^4 \tag{5}$$

where $\beta \in [0,\infty]$ and $\gamma \in [0, 2\pi/3]$.

3 Critical Point Symmetries

The equilibrium shape associated with the GCM Hamiltonian can be obtained by determining the minimum of the energy surface with respect to the geometric variables β and γ, *i.e.* where the first derivative vanish.

Since the parameter C_3 controls the steepness of the potential, and therefore, the dynamical fluctuations in γ, it strongly affects the energies of excited intrinsic states. The parameter $C_3 = 0$ gives a γ-flat potential and an increase of C_3 introduces a γ-dependence in the potential with a minimum at $\gamma = 0$. Changing C_3 will indeed induce a γ-unstable to the symmetric rotor transition; it is best to simultaneously vary C_2 and C_4 as well.

The shape transition from vibrator to rotors is achieved by starting from the vibrator limit, lowering C_2 from positive to negative value, increasing C_4 to large positive value, with gradually increasing C_3 (lowering C_2 from positive to negative value, introducing a large positive C_4 and a positive C_3).

4 Numerical Results Applied to Lanthanide and Actinide chains

The first nucleus to be identified as exhibiting transition from spherical to axially deformed shapes was ^{152}Sm [18], followed by ^{150}Nd [24]. Further work on ^{152}Sm [25] and ^{150}Nd

[25, 26] reinforced this conclusion. In our calculation we will examine and systematically study the lanthanide $^{144-154}$Nd and $^{146-156}$Sm, isotopes and actinide $^{224-234}$Th and $^{230-238}$U isotopes because of the richness of available experimental data indicating a transition of nuclear shapes from spherical to deformed form. The optimized model parameters for each

Table 1: The GCM parameters by (MeV) as derived in fitting procedure used in the calculation.

Nucleus	C_2	C_3	C_4
^{144}Nd	12.46084	1.06407	−26.29034
^{146}Nd	7.98904	8.46249	−5.34827
^{148}Nd	−19.84450	41.41216	105.62500
^{150}Nd	−56.19267	83.37305	248.96600
^{152}Nd	−73.70551	104.57310	319.48270
^{154}Nd	−84.13947	118.02790	362.71460
^{146}Sm	14.49576	1.27688	−30.52593
^{148}Sm	8.89235	9.87290	−5.28215
^{150}Sm	−23.19850	47.32818	121.87500
^{152}Sm	−63.80397	93.79468	281.39990
^{154}Sm	−82.44842	116.19230	356.21830
^{156}Sm	−93.05583	129.83070	400.10950
^{224}Th	0.55766	4.96951	6.10300
^{226}Th	−0.11521	6.38937	9.70762
^{228}Th	−0.83906	7.98671	13.68875
^{230}Th	−1.63871	9.76153	18.10188
^{232}Th	−2.59264	11.71384	23.12250
^{230}U	−1.67560	9.76153	18.18437
^{232}U	−2.63289	11.71384	23.21250
^{234}U	−3.77666	13.84363	28.92012
^{236}U	−4.90299	16.15090	34.85125
^{238}U	−6.23928	18.63565	41.51437

nucleus was adjusted by fitting procedure using a computer simulated search program in order to describe the gradual change in the structure as neutron number varied and to reproduce the properties of the selected reliable state of positive parity excitation ($2_1^+, 4_1^+, 6_1^+, 8_1^+, 0_2^+, 2_3^+, 4_3^+, 2_2^+, 3_1^+$, and 4_2^+) and the two neutron separation energies of all isotopes in each isotopic chain. The resulting parameters are listed explicitly in Table 1. For the isotopic chains investigated here, the collective properties are illustrated by representing the calculated PES describing all deformation effects of the nucleus. We investigated the change of nuclear structure within these chains as illustrated in Figures 1-4. The PES's versus the deformation parameter β for lanthanide and actinide isotopic chains of nuclei evolving from spherical to axially symmetric well deformed nuclei. We remark that for all mentioned nuclei, the PES is not flat, exhibiting a deeper minimum in the prolate ($\beta > 0$) region and a shallower minimum in the oblate ($\beta < 0$) region. Relatively flat PES occur for the N = 86 nuclei ^{146}Nd and ^{148}Sm. A first order shape phase transition with change in number of neutrons when moving from the lighter to heavier isotopes, *i.e* U(5) - SU(3) transitional region are observed.

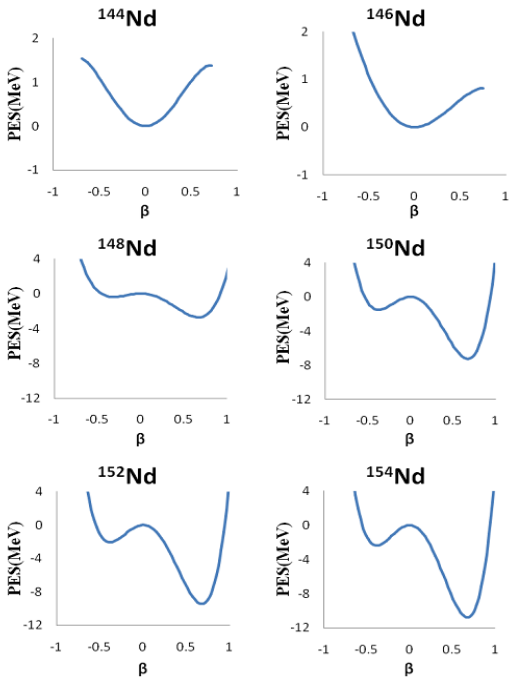

Fig. 1: PES calculated with GCM as a function of the shape parameter β for shape phase transition from spherical to prolate deformed for Neodymium isotope chain $^{144-154}$Nd.

The present results for $^{146-156}$Sm is in good agreement with Nilsson-Strutinsky (BCS)-calculations [26]. However, the existence of a bump in the PES is related to the success of the confined β-soft (BCS) rotor model, employing an infinite square well potential displaced from zero, as well as to the relevance of Davidson potentials [27, 28]. It also is related to the significant five-dimensional centrifugal effect [28, 29]. The actinide $^{228-234}$Th and $^{234-238}$U are all well-deformed rotors with energy ratio $E(4_1^+)/E(2_1^+)$ close to (3.3).

5 Conclusion

A simple approach of the GCM is discussed which reproduces the basic features of the three limits of the nuclear structure: spherical vibrator, axially symmetric rotor and γ-soft rotor, as well as the three phase shape transition regions linking them. The Hamiltonian is expressed as a series expansion in terms of surface deformation coordinates and a conjugate momentum. We considered only the lowest kinetic energy terms, so that the eigen problem for our Hamiltonian reduces to Schrodinger equation in five dimensional spaces. All calculations are performed for reference value of the common mass parameter, only a maximum of three parameters of the truncated form of GCM potential instead of the six are

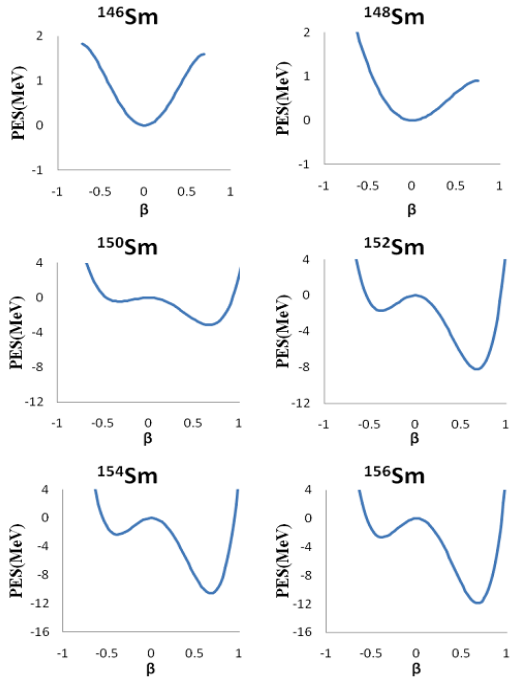

Fig. 2: PES calculated with GCM as a function of the shape parameter β for shape phase transition from spherical to prolate deformed for Samarium isotope chain $^{146-156}$Sm.

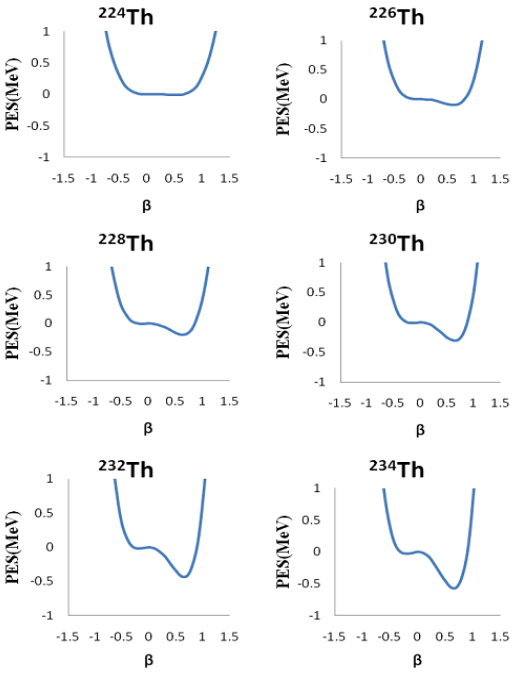

Fig. 3: PES calculated with GCM as a function of the shape parameter β for shape phase transition from spherical to prolate deformed for Thorium isotope chain $^{224-234}$Th.

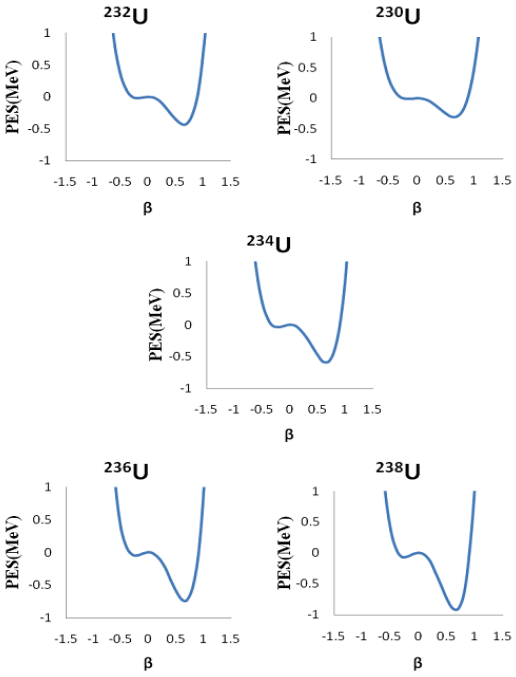

Fig. 4: PES calculated with GCM as a function of the shape parameter β for shape phase transition from spherical to prolate deformed for Uranium isotope chain $^{230-238}$U.

used. The parameter values for the description of a particular nucleus have been found through automated fitting of the nuclear energy levels.

The systematics of shape transitions versus neutron number is studied by the GCM. The capabilities of the model and the illustrative way of representing the collective properties by potential energy surfaces are demonstrated. For neutron number N = 90, the nucleus has a substantial static deformation, but for N = 80 the nucleus is soft or transitional and cannot be described as deformed.

Submitted on December 13, 2012 / Accepted on January 7, 2013

References

1. Khalaf A.M. and Awwad T.M. A Theoretical Description of U(5)-SU(3) Nuclear Shape Transitions in the Interacting Boson Model. *Progress in Physics*, 2013, v. 1, 7–11.
2. Iachello F. and Arima A. The Interacting Boson Model. Cambridge University Press, Cambridge, England, 1987.
3. Frank A. and VanIsacker P. Algebraic Methods in Molecular and Nuclear Structure Physics. Wiley, New York, 1994.
4. Casten R.F. Nuclear Structure from a Simple Prespective. Oxford University, Oxford, 1990.
5. José M.A. et al. Two-Level Interacting Boson Models Beyond the Mean Field. *Physical Review*, 2007, v. C75, 014301–014310.
6. Iachello F. and Zamfir N.V. Quantum Phase Transitions in Mesoscopic Systems. *Physical Review Letters*, 2004, v. 92 (3), 212501–212505.
7. Cejnar P., Heinze S. and Dobes J. Thermodynamic Analogy for Quantum Phase Transitions at Zero Temperature. *Physical Review*, 2005, v. C71, 011304R–011309R.

8. Rowe D.J., Turner P.S. and Rosensteel G. Scaling Properties and Asymptotic Spectra of Finite Models of Phase Transitions as They Approach Macroscopic Limits. *Physical Review Letters*, 2004, v. 93, 232502–232505.
9. Rowe D.J. Quasi Dynamical Symmetry in an Interacting Boson Model Phase Transition. *Physical Review Letters*, 2004, v. 93, 122502–122505.
10. Liu Y.X., Mu L.Z. and Wei H. Approach to The Rotation Driven Vibrational to Axially Rotational Shape Phase Transition Along The Yrast Line of a Nucleus. *Physics Letters*, 2006, v. B633, 49–53.
11. Zhang Y., Hau Z. and Liu Y.X. Distinguishing a First Order From a Second Order Nuclear Shape Phase Transition in The Interacting Boson Model. *Physical Review*, 2007, v. C76, 011305R–011308R.
12. Arios J.M., Dukelsky J. and Garcia-Ramos J.E. Quantum Phase Transitions in the Interacting Boson Model: Integrability, Level Repulsion and Level Crossing. *Physical Review Letters*, 2003, v. 91, 162502–162504.
13. Garcia-Ramos J.E. et al. Two-Neutron Separation Energies, Binding Energies and Phase Transitions in The Interacting Boson Model. *Nuclear Physics*, 2001, v. A688, 735–754.
14. Liu M.L. Nuclear Shape-Phase Diagrams. *Physical Review*, 2007, v. C76, 054304–054307.
15. Heyde K. et al. Phase Transitions Versus Shape Coexistence. *Physical Review*, 2004, v. C69, 054304–054309.
16. Iachello F. Dynamic Symmetries at The Critical Point. *Physical Review Letters*, 2000, v. 85, 3580–3583.
17. Iachello F. Analytic Prescription of Critical Point Nuclei in a Spherical Axially Deformed Shape Phase Transtion. *Physical Review Letters*, 2001, v. 87, 052502–052506.
18. Casten R.F. and Zamfir N.V. Evidence for a Possible E(5) Symmetry in ^{134}Ba. *Physical Review Letters*, 2000, v. 85, 3584–3586. Casten R.F. and Zamfir N.V. Empirical Realization of a Critical Point Description in Atomic Nuclei. *Physical Review Letters*, 2001, v. 87, 052503–052507.
19. McCutchan E.A., Zamfir N.V. and Casten R.F. Mapping The Interacting Boson Approximation Symmetry Triangle: New trajectories of Structural Evolution of Rare-Earth Nuclei. *Physical Review*, 2004, v. C69, 064306–064314. Bridging The Gap Between X(5) and The Interacting Boson Model. *Physical Review*, 2005, v. C71, 034309–034314.
20. McCutchan E.M., Bonatsos D. and Zamfir N.V. Connecting The X(5)-?2, X(5)-?4, and X(3) Models to the Shape/Phase-Transition Region of The Interacting Boson Model. *Physical Review*, 2006, v. C74, 034306–034316.
21. Dieperink A.E.L., Scholten O. On Shapes and Shape Phase Transitions in The Interacting Boson Model. *Nuclear Physics*, 1980, v. A346, 125–138.
22. Bohr A. and Mottelson. Nuclear Structure v. II. Benjamin, New York, 1975.
23. Troltenier D., Hess P.O. and Maruhn J. Computational Nuclear Physics, Vol. I, Nuclear Structure, Langanke eds.K., Maruhn J. and Koonin S.E. Springer, Berlin, Heidelberg, New York, 1991.
24. Krücken R. et al. B(E2) Values in ^{150}Nd and The Critical Point Symmetry X(5). *Physical Reviw Letters*, 2002, v. 88, 232501–232501.
25. Bijker R. et al. Test of X(5) for the γ Degree of Freedom. *Physical Reviw*, 2003, v. C68, 064304–064307 and Erratum: Test of X(5) for the γ Degree of Freedom. *Physical Review*, 2004, v. C69, 059901–059905.
26. Zhang D.L. and Zhao H.Y.C. Lifetime measurements in the γ band of ^{160}Er nuclei. *Physical Letters*, 2002, v. 19, 779–782.
27. Bonatsos D. et al. E(5) and X(5) Critical Point Symmetries Obtained From Davidson Potentials Through a Variational Procedure. *Physical Review*, 2004, v. C70, 024305–024314.
28. Caprio M.A. Effects of $\beta - \gamma$ coupling in transitional nuclei and the validity of the approximate separation of variables. *Physical Review*, 2005, v. C72, 054323–054333.
29. Rowe D.J. and Turner P.S. The Algebraic Collective Model. *Nuclear Physics*, 2005, v. A753, 94–105.

Quantum Uncertainty and Fundamental Interactions

Sebastiano Tosto
Italy. Email: stosto@inwind.it, stosto44@gmail.com

The paper proposes a simplified theoretical approach to infer some essential concepts on the fundamental interactions between charged particles and their relative strengths at comparable energies by exploiting the quantum uncertainty only. The worth of the present approach relies on the way of obtaining the results, rather than on the results themselves: concepts today acknowledged as fingerprints of the electroweak and strong interactions appear indeed rooted in the same theoretical frame including also the basic principles of special and general relativity along with the gravity force.

1 Introduction

The state of a classical particle is specified by its coordinates and momentum; the dynamical variables x, p_x, y, p_y, z, p_z, assumed known at any time, define the 6-dimensional space usually called "phase space". Knowing the state of a particle means determining these six quantities that describe its motion and energy. Since the state of a classical system is identified by the distribution of corresponding points in the phase space, any finite volume $V_{ps} = (\delta x\delta y\delta z)(\delta p_x\delta p_y\delta p_z)$ should seemingly contain an infinite number of states. Because of the uncertainty principle, however, these six quantities are not simultaneously known; the impossibility of defining the corresponding points in the phase space compels instead introducing a lower limit to the volume of phase space physically significant. Since such an elementary volume has size $V^o_{ps} = (dxdydz)(dp_xdp_ydp_z) = \hbar^3$, any finite volume V_{ps} enclosing measurable combinations of coordinates and conjugate momenta consists of a finite number V_{ps}/V^o_{ps} of elementary volumes. The quantum uncertainty was inferred by W. Heisenberg as a consequence of the operator formalism of wave mechanics, on which relies the quantum theory: the wave function $\psi = \psi(x, t)$ replaces the lack of definable quantum values of x concurrently associable to the conjugate p_x. However most physicists believe unsatisfactory a theory based on the wave function ψ without direct physical meaning [1]; indeed $\psi\psi^*$ only has the statistical meaning of probability density and contains the maximum information obtainable about a physical system. The wave function characterizes a pure state, represented by a single "ket" vector to which corresponds a well defined eigenvalue, whereas in general a particle is found in a mixture of states; so the result of a measurement on a quantum state represents a probability distribution of finding the particle in a given volume of phase space. The density matrix is the mathematical tool to describe mixed quantum states by means of a distribution function of coordinates and momenta. Owing to the statistical character of the knowledge we can afford in the quantum world, the Wigner function $W(x, p)$ [2] aims to represent a quantum state in terms of a joint probability distribution involving both coordinates and momenta, in formal analogy with the classical statistics; the former is therefore a correction to the latter. The quantum x and p distributions are appropriately described by the respective marginal distributions $\int_{-\infty}^{+\infty} W(x, p)dp$ and $\int_{-\infty}^{+\infty} W(x, p)dx$ under the normalization condition $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} W(x, p)dpdx = 1$, whereas the expectation value for any operator function is weighed by $W(x, p)$ as $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} W(x, p) f(x, p)dpdx$. Other relevant features of $W(x, p)$, well known [3], are omitted here for brevity. Also the Wigner function, however, although providing significant information about the quantum states, presents conceptual difficulties: it is not a real probability distribution in the classical sense, it is a quasi-probability that can even take negative values; moreover it can represent the average value of an observable but not, in general, also its higher power moments.

To bypass both these difficulties inevitably inherent the wave formalism, the present theoretical model implements an approach conceptually different: it exploits directly the statistical formulation of quantum uncertainty, which therefore becomes itself a fundamental assumption of the model and reads in one space dimension

$$\Delta x\Delta p_x = n\hbar = \Delta t\Delta\varepsilon. \qquad (1,1)$$

This set of $2n$ equation disregards since the beginning the local dynamical variables of the particles forming the quantum system and simply counts its number n of allowed states. Are therefore required the following positions

$$x_i \to \Delta x_i, \qquad t \to \Delta t, \qquad i = 1..3. \qquad (1,2)$$

No hypotheses are made about the uncertainty ranges, which are by definition unknown, unknowable and arbitrary. In quantum mechanics the square complex wave function of space and time variables contains the maximum information about a quantum system, which has therefore probabilistic character. The present model intends instead starting from a *minimal* information about any quantum system, still based on the failure of the physical concept of points definable in the quantum phase space but trusting on the idea that a minimum information is consistent with the maximum generality: despite the knowledge of one dynamical variable only is in principle allowed even in the quantum world, the present

model disregards "a priori" the local values of both conjugate dynamical variables. This means renouncing even to the concept of probability density provided by the wave function of a particle, while also disregarding the related concept of wave packet to describe its propagation; in the present model it is only possible to say that if the particle moves during a time range Δt throughout its uncertainty range Δx, then its average velocity component is $v_x = \Delta x/\Delta t$ regardless of any local feature of its actual delocalization motion. So eqs (1,1) require by definition $\Delta\varepsilon = v_x \Delta p_x$. In fact the positions (1,2) ignore both local dynamical variables, not as a sort of approximation to simplify some calculation but conceptually and since the early formulation of any quantum problem; accordingly, the delocalization of a quantum particle in its uncertainty range is conceived in its most agnostic form, i.e. waiving any kind of information about its position and motion. Thus, regarded in this way, eqs (1,1) exclude the concept itself of probability density and contextually also the definition of Wigner function linking the Schrodinger equation to the marginal distributions in the phase space; both equations are bypassed along with the concept of wave equation itself. Eqs (1,1) merely list the eigenvalues of pure states, indeed they are a set of equations corresponding to the respective values of n; so they also skip the probability with which in a mixed state each eigenvalue could be measured. Despite waiving themselves the concept of probability density through the positions (1,2), eqs (1,1) enable however also this kind of probabilistic information; it is essential indeed to mention that the wave formalism is obtainable as a corollary of eqs (1,1) [4], which means that all considerations previously introduced are in fact comprised also in the present theoretical model: one infers first from eqs (1,1) the operator formalism and then proceeds as usual. In this way the wave formalism, with its conceptual weakness, loses its rank of fundamental root of our knowledge about the quantum world, becoming indeed a mere by-product of eqs (1,1); yet, even so it still represents an added value to the physical information by introducing the concept of probability density that partially overcomes the total agnosticism of eqs (1,1).

What however about the chance of formulating any physical problem exploiting directly the eqs (1,1) only? Is legitimate the belief that the equations enclosing conceptually the wave formalism as a corollary also enclose the inherent physical information. The question that arises at this point concerns just the real chance of obtaining physical information once abandoning the typical ideas and mathematical tools of wave mechanics: is really redundant the concept of probability density? Several papers have demonstrated the effectiveness of this alternative approach, e.g. [5,6]; moreover, without the need of hypotheses on n and on the uncertainty ranges defined by eqs (1,1), the paper [7] has shown the possibility of extending the mere quantum horizon of these equations, initially concerned, also to the special and general relativity. The positions (1,2) compel focusing the attention on the uncertainty ranges and related numbers of states, i.e. on the phase space, rather than on the specific coordinates of the particles concerned by the particular physical problem. In fact, the local dynamical variables are conceptually disregarded since the beginning in the present model. Put for instance $\Delta x = x - x_o$: if either boundary coordinate, say x_o, is defined by the origin of the coordinate system R, then it determines the position of Δx in R; the other boundary coordinate x determines its size. The crucial point is that both x_o and x are arbitrary, unknown and unknowable by fundamental assumption; the reference system R is therefore "a priori" arbitrary, unspecified and unspecifiable as well, whence the equivalence of all reference systems whenever implementing the positions (1,2) to describe the quantum world. Otherwise stated, eqs (1,1) do not specify any particular reference system because analogous considerations hold for all uncertainty ranges they introduce. Moreover n is itself arbitrary as well; it merely symbolizes a sequence of numbers of allowed states, not some specific value in particular. Let therefore eqs (1,1) be defined in any R and rewrite them as $\Delta x' \Delta p'_x = n' = \Delta\varepsilon' \Delta t'$ in any R': it is self-evident that actually these equations are indistinguishable because n and n' do so as well. Whatever a specific value of n might be in R, any change to n' e.g. because of the Lorentz transformations of the ranges is physically irrelevant: it means replacing an arbitrary integer in the former set with another integer of the latter set. In effect, two examples of calculation reported below highlight that modifying the range sizes from primed to unprimed values does not affect any result, in agreement with their postulated arbitrariness: no range size is expected to appear in the quantum eigenvalues. Hence the eqs (1,1) have general character, regardless of any particular reference system to be appropriately specified; this holds also if R and R' are inertial and non-inertial, since no hypothesis has been assumed about them [7]. On the one hand this entails obtaining the indistinguishability of identical particles as a corollary, regardless of which particle in a set could be that actually delocalized in a given uncertainty range; indeed no particle is specifically concerned "a priori". On the other hand it also entails that the properties of motion of the particle, and thus the marginal distributions of its dynamical variables, are disregarded by assumption and skipped by consequence when formulating any physical problem. To better understand the following of the paper, these remarks are now exemplified examining shortly the non-relativistic quantum angular momentum $\mathbf{M}$, on the one side to highlight how to exploit the positions (1,2) and on the other side to show why the minimal information accessible through eqs (1,1) is in fact just that available through the usual operator formalism of wave mechanics.

Consider the classical component $M_w = \mathbf{r} \times \mathbf{p} \cdot \mathbf{w}$ of $\mathbf{M}$ along an arbitrary direction defined by the unit vector $\mathbf{w}$, being $\mathbf{r}$ the radial distance of any particle from the origin of an arbitrary reference system and its momentum. The positions (1,2) compel $\mathbf{r} \to \Delta\mathbf{r}$ and $\mathbf{p} \to \Delta\mathbf{p}$ and enable the

number l of states to be calculated only considering the total ranges $\Delta\mathbf{r}$ and $\Delta\mathbf{p}$ of distances and momenta physically allowed to the particle, about which no hypothesis is necessary; let us show that the random local values $\mathbf{r}$ and $\mathbf{p}$ themselves have instead no physical interest. So $M_w = (\Delta\mathbf{r} \times \Delta\mathbf{p}) \cdot \mathbf{w} = (\mathbf{w} \times \Delta\mathbf{r}) \cdot \Delta\mathbf{p}$, i.e. $M_w = \Delta\mathbf{W} \cdot \Delta\mathbf{p}$, where $\Delta\mathbf{W} = \mathbf{w} \times \Delta\mathbf{r}$. If and $\Delta\mathbf{W}$ are orthogonal, then $M_w = 0$; else, rewriting $\Delta\mathbf{W} \cdot \Delta\mathbf{p}$ as $(\Delta\mathbf{p} \cdot \Delta\mathbf{W}/\Delta W)\,\Delta W$ with $\Delta W = |\Delta\mathbf{W}|$, the component $\pm\Delta p_W = \Delta\mathbf{p} \cdot \Delta\mathbf{W}/\Delta W$ of $\Delta\mathbf{p}$ along $\Delta\mathbf{W}$ yields $M_w = \pm\Delta W \Delta p_W$.

Thus, according to eqs (1,1), $M_w = \pm l\hbar$, being l the usual notation for the number of states of the angular momentum. As expected, M_w is a multi-valued function because of the uncertainties initially postulated for $\mathbf{r}$ and $\mathbf{p}$. One component of $\mathbf{M}$ only, e.g. along the z-axis, is knowable; repeating the same approach for the y and x components would trivially mean changing $\mathbf{w}$. Just this conclusion suggests that the average values $< M_x^2 >$, $< M_y^2 >$ and $< M_z^2 >$ should be equal; so the quantity of physical interest to describe the properties of quantum angular momentum is l, as a function of which M^2 is indeed inferred as well. The components averaged over the possible states summing $(l\hbar)^2$ from $-L$ to $+L$, where L is an arbitrary maximum value of l, yield $< M_i^2 >= \sum_{l_i=-L}^{l_i=L} (\hbar l)^2/(2L{+}1)$ and thus $M^2 = \sum_{i=1}^{3} < M_i^2 >= L(L{+}1)\hbar^2$.

The physical definition of angular momentum is enough to find quantum results completely analogous to that of the wave mechanics even disregarding any local detail about the angular motion. This result has been reminded here as it introduces several significant considerations useful in the following: (i) eqs (1,1) and the positions (1,2) plug the classical physics into the quantum world; (ii) no hypothesis is necessary about the motion of the particle nor about its wave/matter nature to infer the quantum result; (iii) trivial algebraic manipulations replace the solution of the pertinent wave equation; (iv) the result inferred through eqs (1,1) only is consistent with that of the wave mechanics; (v) the local distance between the particles concerned in the angular motion does not play any role in determining l; (vi) the number of allowed states plays actually the role of angular quantum number of the operator formalism of wave mechanics; (vii) the amount of information accessible for the angular momentum is not complete like that of the classical physics, but identical to that of the wave formalism; (viii) eqs (1,1) rule out "a priori" any chance of hidden variables hypothetically encodable in the wave function, i.e. local values of any kind that could in principle enhance our knowledge about M_w and M^2 to obtain a more complete description of the angular quantum system; (ix) the eigenvalues, i.e. the physical observables, are actually properties of the phase space rather than properties of specific particles, whence the indistinguishability of identical particles here inferred as a corollary of eqs (1,1); (x) the numbers of states are here simply counted; (xi) the positions (1,2) are consistent with the concept of classical coordinate in the limit case $\Delta x \to 0$, which means that the random local variable $x_o \le x \le x_1$ tends to a classical local value uniquely and exactly defined; (xii) the total arbitrariness of the boundary values of the ranges is necessary to ensure that any local value is allowed for the corresponding classical variables; (xiii) the range sizes do not play any role in determining the eigenvalues of angular momentum, their conceptual reality, i.e. the total uncertainty about both conjugate dynamical variables of a quantum particle, is the unique hypothesis of the present model. The same holds of course for any other uncertainty range.

These ideas have been extended and checked in the papers [5,6] also for more complex quantum systems like hydrogenlike and many electron atoms/ions and diatomic molecules; also these papers allowed concluding that eqs (1,1) efficiently replace the standard approach of wave mechanics, without requiring the concept of probability density and thus without need of calculating marginal distributions in the phase space through the Wigner functions. In these papers the interaction is described via the Coulomb potential energy between charged particles; in other words, one assumes already known the Coulomb law to calculate for instance the energy levels of hydrogenlike atoms. This point is easily highlighted considering for simplicity the non-relativistic hydrogenlike energy levels; also this topic, already introduced in [5], is reported here for completeness.

Assuming the origin O of an arbitrary reference system R on the nucleus, the classical energy is $\varepsilon = p^2/2m - Ze^2/r$ being m the electron mass. Since $p^2 = p_r^2 + M^2/r^2$, the positions (1,2) $p_r \to \Delta p_r$ and $r \to \Delta r$ yield $\varepsilon = \Delta p_r^2/2m + M^2/2m\Delta r^2 - Ze^2/\Delta r$. Two numbers of states, i.e. two quantum numbers, are expected because of the radial and angular uncertainties. Eqs (1,1) and the previous result yield $\varepsilon = n^2\hbar^2/2m\Delta r^2 + l(l+1)\hbar^2/2m\Delta r^2 - Ze^2/\Delta r$ that reads $\varepsilon = \varepsilon_o + l(l+1)\hbar^2/2m\Delta r^2 - E_o/n^2$ with $E_o = Z^2e^4m/2\hbar^2$ and $\varepsilon_o = (n\hbar/\Delta r - Ze^2m/n\hbar)^2/2m$. Minimize ε putting $\varepsilon_o = 0$, which yields $\Delta r = n^2\hbar^2/Ze^2m$ and $\varepsilon = [l(l+1)/n^2 - 1]E_o/n^2$; so $l \le n-1$ in order to get $\varepsilon < 0$, i.e. a bound state. Putting thus $n = n_o + l + 1$ one finds the electron energy levels $\varepsilon_{el} = -E_o/(n_o + l + 1)^2$ and the rotational energy $\varepsilon_{rot} = l(l+1)E_o/n^4$ of the atom as a whole around O. Hold also here all considerations introduced for the angular momentum, in particular it appears that the range sizes do not play any role in determining the energy levels. The physical meaning of Δr, related to the early Bohr radius, appears noting that

$$\varepsilon_{el} = -\frac{E_o}{n^2} = -\frac{Ze^2}{2\Delta r}, \quad \Delta r = \frac{n^2\hbar^2}{Ze^2m}, \quad E_o = \frac{Z^2e^4m}{2\hbar^2}, \quad (1,3)$$

i.e. ε_{el} is due to charges of opposite sign delocalized within a diametric distance $2\Delta r$ apart. As previously stated, nucleus and electron share a unique uncertainty radial range: in general, the greater m, the closer its delocalization extent around the nucleus. Also note that n and l are still properties of the phase space, but now they describe the whole quantum sys-

tem "nucleus + electron" rather than the nucleus and the electron separately. Since the first eq (1,3) does not depend explicitly on the kind of particles forming the concerned hydrogenlike atom, m or the reduced mass are actually hidden into Δr; it is possible to link ε_{el} to the known condition $n\lambda = 2\pi\Delta r$, according which an integer number of steady electron wavelengths is defined along a circumference of radius Δr. For such electron waves one finds

$$\varepsilon_{el} = -\frac{\pi Z e^2}{n\lambda} = -\frac{\alpha}{n}\frac{Z p_\lambda c}{2}, \quad p_\lambda = \frac{h}{\lambda}, \quad \alpha = \frac{e^2}{\hbar c}. \qquad (1,4)$$

Note that introducing α to express the quantum energy levels compels defining the De Broglie momentum. Even in this form ε_{el} is still related to the reduced mass of the system, which can be introduced via the momentum p_λ; thus eq (1,4) holds in general for any system of charges. Moreover, the factor $Z/2$ apart, appears interesting that the energy levels of the system ε_{el} are linked to the kinetic energy $p_\lambda c$ of the running electron wave circulating along the circumference of radius Δr via the coefficient α/n. On the one hand, this result emphasizes the electromagnetic character of the interaction between electron and nucleus; on the other hand, the key role of the quantum uncertainty in determining the allowed energy levels of eqs (1,3) also evidences the kind of interaction itself. The more general question that arises at this point is therefore: do eqs (1,1) provide themselves any hint also about the physical essence of the fundamental interactions? The standard model [8-11] provides a satisfactory description of the fundamental forces of nature. So the present paper does not aim to replicate the electro-weak model or the chromodynamics, which indeed would be useless and unexciting; nevertheless seems useful to propose a simplified approach aimed to show (i) that the fundamental interactions are inferable from eqs (1,1) only and (ii) that exists a unique conceptual root common to all fundamental interactions. This task is in effect particularly valuable because the present model has already accounted for the gravity force [7] and for the basic principles of special and general relativity.

The purpose of the paper is to examine the ability of eqs (1,1) to describe also other kinds of possible interactions and their relative strengths at comparable energies; it will be also shown that further information is obtained about the vector bosons associated with the respective kinds of interactions. Therefore the worth of the present paper rests mostly on the chance of finding concepts today known as fingerprints of the electroweak and strong interactions in the frame of a unique logical scheme based on the quantum uncertainty and including the relativity. The paper [7] has somewhat concerned the electromagnetic interactions, while also showing that all concepts of quantum wave formalism are indeed obtained through the present approach. Here we concern in particular the weak and strong interactions between nuclear and subnuclear particles. The next sections will describe the possible features of these interactions.

2 Physical background of the interactions

Let us show that the concept of interaction relies in the frame of the present model entirely on eqs (1,1). Consider first an isolated particle of mass m and momentum component p_x^∞ free to move in an ideal infinite range. When confined in a time-space uncertainty range Δx, however, its energy changes by an amount $\Delta\varepsilon$ given by

$$\Delta p_x^2/2m = (n\hbar)^2/2m\Delta x^2, \qquad \Delta p_x = p_x^{conf} - p_x^\infty;$$

i.e. Δp_x is by definition the range including any change of local momentum component p_x occurring when the free particle turns from a non-confined to a confined state within Δx.

Since no process occurs instantaneously in nature, let Δt be the confinement time range corresponding to Δp_x: to the confinement process corresponds thus the arising of a force field whose component $\Delta F_x = \Delta p_x/\Delta t = F_x^{conf} - F_x^\infty$ is related to $\Delta\varepsilon$, being clearly $\Delta F_x = \Delta\varepsilon/\Delta x = \Delta p_x^2/2m\Delta x^3$. By definition ΔF_x includes any random $F_x^\infty \le F_x \le F_x^{conf}$: in the present model the local dynamical variables are replaced by corresponding ranges of values, so the classical force F_x at the local coordinate x is replaced by a range of possible forces active within Δx. Actually the result $\Delta p_x/\Delta t = \Delta\varepsilon/\Delta x$ could have been inferred directly from eqs (1,1) without need of any remark; yet these considerations highlight that a force field in a space time uncertainty range is the only information available on the particle once accepting the eqs (1,1) as the unique assumption of the model.

Clearly, once concerning one particle only, energy and force component cannot be related to any form of interaction; rather both have mere quantum origin. Also, $\Delta\varepsilon$ and ΔF_x tend obviously to zero for $\Delta x \to \infty$; hence if p_x^∞ changes to p_x^{conf} concurrently with the arising of a force component acting on the particle, then p_x^∞ must be constant by definition as it represents the momentum of the particle before its confinement driven perturbation. This again appears from the standpoint of eqs (1,1): $\Delta x \to \infty$ requires $\Delta p_x \to 0$ for any finite number of states regardless of Δt. Since an uncertainty range infinitely small tends to a unique classical value of its corresponding quantum random variable and since this holds regardless of Δt, then the limit value must be a constant: so $p_x^\infty = const$ corresponds by necessity to $F_x^\infty = 0$.

Despite the present model allows reasoning on ΔF_x only, a first corollary is the inertia principle that holds for a lonely particle in an infinite space time delocalization range. Other interesting consequences follow for any finite $\Delta x = x_2 - x_1$: the notation emphasizes that instead of considering the particle initially in an infinite unconfined range, we are now interested to describe its behavior in a confined state, e.g. in the presence of two infinite potential walls Δx apart. Clearly this means introducing the corresponding $\Delta p_x = p_2^{conf} - p_1^{conf}$: again the eqs (1,1) compel writing $\Delta\varepsilon/\Delta x = \Delta p_x^2/2m\Delta x^3$ when p_x^∞ has turned into a local $p_1^{conf} \le p_x \le p_2^{conf}$, which entails once more $\Delta F_x = \Delta p_x/\Delta t$ within Δx. These ideas are

now extended to the interaction forces. Rewrite first the force field component $\Delta\varepsilon/\Delta x = \Delta p_x^2/2m\Delta x^3$ of a particle confined within Δx as follows

$$\Delta F_x = \frac{\hbar^2}{2}\frac{n}{m}\frac{n}{V}, \qquad V = \Delta x^3. \tag{2,1}$$

Even the one-dimensional case defines the delocalization volume V because, being Δx, Δy and Δz arbitrary, any value allowed to $\Delta x\Delta y\Delta z$ is also allowed to Δx^3. Is crucial the fact that the range of each force component is proportional to n/m, number of allowed states per unit mass, times n/V, number of allowed states per unit delocalization volume. Consider now two free particles a and b in their own uncertainty ranges Δx_a and Δx_b; hold separately for them the relationships $\Delta\varepsilon_a = (n_a\hbar)^2/2m_a\Delta x_{xa}^2$ and $\Delta\varepsilon_b = (n_b\hbar)^2/2m_b\Delta x_{xb}^2$. These particles are non-interacting, as their n_a and n_b are assumed independent each other like Δx_a and Δx_b themselves; nothing in these equations accounts for the most typical and obvious consequence of any kind of interaction, i.e. some relationship between their allowed states or between their delocalization ranges. Two free particles do not share by definition any kind of link, any possible coincidence of allowed states would be accidental and transient only. Consider now their possible interaction; a reasonable chance of linking their allowed states is to assume, for instance, that the particles share the same uncertainty range. If Δx is unique for both particles, then their allowed states must be somehow linked because of eqs (1,1); in other words, even being still $n_a \neq n_b$, the random values of local momentum components p_{xa} and p_{xb} are subjected to the constrain $n_a/\Delta p_{xa} = n_b/\Delta p_{xb} = \Delta x/\hbar$. Note for instance that Δr of eq (1,3) includes by definition all possible distances between electron and nucleus, which implicitly means that both particles share the same uncertainty range where the interaction occurs; so n and l characterizing the electron energy levels of the hydrogenlike system result from the change of the early quantum numbers, e.g. n^{free} and $l^{free} = 0$, owned by each particle independently of the other before interaction. In this respect two relevant points are: (i) the interaction driven change δn of the number n of states and (ii) the physical meaning of the related $\delta[(n/m)(n/V)]$.

As concerns the point (i), consider $\Delta\varepsilon\Delta t = n\hbar$ in an arbitrary reference system R and let n be allowed to change from any initial value n_1 to any successive value n_2 during a fixed time range Δt; whatever n_1 and n_2 might be, this is admissible because Δt is arbitrary. The notation emphasizes that a given value of $\delta n = n_2 - n_1$ is obtainable regardless of the initial value n_1 because n_2 is arbitrary; so $\delta n = 1, 2, ..$ anyway, regardless of the specific value of n_1. Calculate next the change $\delta\Delta\varepsilon$ of $\Delta\varepsilon$ as a function of δn during Δt, which reads now $(\Delta\varepsilon_{n_2} - \Delta\varepsilon_{n_1})/\Delta\varepsilon_{n_1} = \delta n/n_1$ with obvious meaning of symbols. Note that in general the series expansion of $\log(\Delta\varepsilon)$ around $\log(\Delta\varepsilon_{n_1})$ reads

$$\log(\Delta\varepsilon_{n_2}) = \log(\Delta\varepsilon_{n_1}) + \frac{\Delta\varepsilon_{n_2} - \Delta\varepsilon_{n_1}}{\Delta\varepsilon_{n_1}} - \frac{1}{2}\left(\frac{\Delta\varepsilon_{n_2} - \Delta\varepsilon_{n_1}}{\Delta\varepsilon_{n_1}}\right)^2 + ..$$

so that

$$\log\left(\frac{\Delta\varepsilon_{n_1+\delta n}}{\Delta\varepsilon_{n_1}}\right) = \frac{\delta n}{n_1} - \frac{1}{2}\left(\frac{\delta n}{n_1}\right)^2 + \frac{1}{3}\left(\frac{\delta n}{n_1}\right)^3 - ..$$

$$\Delta\varepsilon_{n_1} = \frac{n_1\hbar}{\Delta t}, \quad \delta n = 1, 2, \ldots \tag{2,2}$$

This equation describes the size change of the energy range $\Delta\varepsilon_{n_1}$ as long as the number of allowed states increases with respect to the initial value n_1: so $\Delta\varepsilon_{n_1+\delta n}$ with $\delta n = 1$ describes the first increment of energy range size with respect to $\Delta\varepsilon_{n_1}$, then $\delta n = 2$ the next size increment and so on; in short, eq (2,2) describes how are modified the random local values $\varepsilon_{n_1+\delta n}$ included in $\Delta\varepsilon_{n_1+\delta n}$ at δn progressively increasing. Instead $\Delta\varepsilon_{n_1}$ plays here the role of a fixed reference range with respect to which is calculated $\Delta\varepsilon_{n_1+\delta n}$. For reasons that will be clear in the next section 5, it is mostly interesting to examine the particular case of n_1 such that

$$\Delta\varepsilon_{n_2} - \Delta\varepsilon_{n_1} << \Delta\varepsilon_{n_1}, \qquad \delta n/n_1 << 1. \tag{2,3}$$

Let us truncate thus the series expansion (2,2) at the first order of approximation under the assumption (2,3) and simplify the notation putting $i = \delta n$; one finds (i=1,2,...)

$$n_1 \log\left(\frac{\Lambda_i}{\Lambda}\right) = i, \qquad \Lambda_i = \Delta\varepsilon_{n_1+\delta n}, \qquad \Lambda = \Delta\varepsilon_{n_1}. \tag{2,4}$$

Despite the generality of eqs (2,2), is particularly significant for the purposes of the present paper the case of a quantum system consisting of an arbitrary number of particles, each one delocalized in its own uncertainty range: if these latter are non-interacting, then let the energy of the system be included within the range $\Delta\varepsilon_{n_1}$ and be n_1 its total number of states; if instead all particles are delocalized in the same space-time range, then their interaction changes the energy range of the system to $\Delta\varepsilon_{n_1+\delta n}$ characterized of course by a new number of states $n_2 = n_1 + \delta n$.

As concerns the point (ii), we expect according to eq (2,1) that from $\Delta\varepsilon_a$ and $\Delta\varepsilon_b$ of the two free particles follow because of the interaction the changes $\delta\Delta\varepsilon_a = (\hbar^2/2)\delta(n_a^2/m_a\Delta x_a^2)$ and $\delta\Delta\varepsilon_b = (\hbar^2/2)\delta(n_b^2/m_b\Delta x_b^2)$. The expressions of the corresponding changes of the initial confinement force components $\Delta F_{xa} = \Delta\varepsilon_a/\Delta x_a$ and $\Delta F_{xb} = \Delta\varepsilon_b/\Delta x_b$ from the non-interacting to the interacting state read thus

$$\delta\Delta F_{xa} = (\hbar^2/2)\delta\left[(n_a/m_a)(n_a/V_a)\right]$$

$$\delta\Delta F_{xb} = (\hbar^2/2)\delta\left[(n_b/m_b)(n_b/V_b)\right].$$

These equations agree with the previous idea, i.e. the forces are related to changes of the allowed numbers of states per unit mass and delocalization volumes of the particles a and b: in effect the interaction between two particles consists of forces acting on both of them and requires that the respective numbers of states are affected as well. More precisely

$\delta[(n/m)(n/V)]$ means that are modified during the interaction not only the states allowed to the particles themselves, but also that of the delocalization space surrounding them. Clearly the former are consequences of the latter. In other words, the fact that $\delta(n/m)$ requires explicitly also the concurrent $\delta(n/V)$ compels thinking: (i) that a particle interacts with another particle because it generates a field that propagates outwards through the space volume V and (ii) that just in doing so this field changes the number of states allowed to the other particle; i.e. the changes of number of states of each particle are somehow correlated, as previously stated. Since no event occurs instantaneously in nature, $\delta(n/V)$ requires an appropriate time range to be realized, i.e. the propagation rate is finite in agreement with the existence of an upper limit obliged by eqs (1,1) [7]; in this way the interaction exchanges information about physical features and strength of the related force between particles. The most natural way to acknowledge this way of regarding two interacting particles is to admit that they exchange intermediate virtual particles that propagate, whence $\delta(n/V)$, and carry the necessary information that affects in turn the real particles themselves, whence $\delta(n/m)$; indeed n defining n/V is the same as that defining n/m, i.e. the change $\delta(n/m)$ of states allowed to the particle is actually just that $\delta(n/V)$ of the space around it. Strictly speaking, however, one should say more appropriately space-time, and not simply space: indeed Δx defining V in eq (2,1) is actually $\Delta x = \Delta x(\Delta t)$ because of eqs (1,1) themselves. So the finite time range required for $\delta(n/m)$ to occur is nothing else but the finite time range required to propagate $\delta(n/V)$ and to come back, i.e. to allow exchanging the interaction carriers. Interaction force and propagation of force carriers through V are therefore according to eq (2,1) two basic aspects of the interaction. In principle these carriers could be massive or massless, in which case one expects $(\hbar^2c^2)\delta[(n/\varepsilon)(n/V)]$, but they must have anyway boson character in order that the aforesaid forces affect the allowed states of the interaction partners while minimizing their exchange energy. It has been already demonstrated in [7] that as a consequence of eqs (1,1) integer or half-integer spin particles have a different link to the respective numbers of allowed states: an arbitrary number of the former can be found in a given quantum state, instead one particle only of the latter kind can be found in a given quantum state. Consider a multi-body interaction, where an arbitrary number of force carriers is to be expected: fermion carriers would require a corresponding number of quantum states with energy progressively increasing, whereas a unique ground state allows any number of boson carriers; as it will be shown below, the former case would be incompatible with a unique amount of energy to be transferred between all interacting particles and thus with at a minimum transfer energy. The corpuscles that mediate the fundamental forces of nature are indeed well known in literature as vector bosons, which also suggests the existence of a pertinent boson energy field. An interesting consequence of eq (2,1) comes from the chance of rewriting it as $(m/n\hbar)\Delta F_x = (\hbar/2)(n/V)$. Note that at left hand side appears the ratio $\hbar/m$ having physical dimensions of diffusion coefficient; write therefore $\Delta F_x = D^* n\hbar/2V$ with $D^* = n\hbar/m$. Moreover the fact that the physical dimensions of F/D^* are $mass/(length \times time)$ suggests the position

$$\frac{\Delta F_x}{D^*} = \frac{\hbar}{2}\frac{n}{V} = \frac{du_\omega}{d\omega}, \quad D^* = \frac{n\hbar}{m}, \tag{2,5}$$

having at the moment mere formal meaning: if ω represents a frequency and u_ω an energy density, the physical dimensions of both sides are $energy \times time/volume$. So $\Delta F_x = D^* du_\omega/d\omega$ agrees with the idea that the force field is due to a diffusion-like flux of particles. This appears properly handling $du_\omega/d\omega$: indeed it is possible to write $du_\omega/d\omega = \omega V d\mathrm{C}/dx$ once more via dimensional requirement, being $\mathrm{C} = m/V$ or $\mathrm{C} = \varepsilon/c^2V$ the concentration of massive or massless carriers. Hence $\Delta F_x = \omega V D^* d\mathrm{C}/dx$ i.e. $\Delta F_x = -\omega V J_x$; the minus sign means of course an incoming flux of messenger particles if $J_x > 0$, yet both signs possible for $d\mathrm{C}$ reveal a complex fluctuation driven space distribution of interaction carriers randomly moving forwards and backwards between the real particles. This result is easily understood: in a volume V where are delocalized interacting particles, boson carriers with density C are exchanged at frequency ω according to a Fick-like law that generates the force field ΔF_x; the flow J_x of vector bosons crosses an ideal plane perpendicular to the flow moving at rate $\omega\Delta x$ consistently with an energy $\Delta F_x \Delta x/V$ per unit volume. The diffusion coefficient of the bosons is quantized. In [12] has been demonstrated the quantum nature of the diffusion process and also the link between particle flow and concentration gradient driven Fick's law, as a consequence of which the statistical nature of the entropy also follows; this latter result is further inferred in the next section 7 in an independent way, see eqs (7,7). Eq (2,5) is immediately verifiable considering the cubic volume $V = \Delta x^3$ of space of eq (2,1) filled with photons. Let $\Delta x = \lambda$ be the longest wavelength allowed in V to a steady electromagnetic wave with nodes at the opposite surfaces of the cube, whose side is therefore $\lambda/2$; thus $V = (\lambda/2)^3$, whereas $u_\omega = (\hbar\omega/2)/V$ is the corresponding zero point energy density of the oscillating electromagnetic field. So, with $\lambda = c/\nu$ one finds $du_\omega = 4n(\nu/c)^3\hbar d\omega$; since by definition $\hbar d\omega = h d\nu$, and thus $du_\omega = (2\pi)^{-1} du_\nu$, this result reads $du_\nu = (8\pi(\nu/c)^3 h d\nu)n$. In section 7 it will be shown that the number of states n allowed to the photons trapped within the cube is given by $(\exp(h\nu/kT) - 1)^{-1}$, whence the well known result

$$\frac{du_\nu}{d\nu} = \frac{8\pi h\nu^3}{c^3}n, \qquad n = \frac{1}{\exp(h\nu/kT) - 1}. \tag{2,6}$$

It is interesting the fact that the black body law comes immediately from the same idea that shows the existence of messenger bosons mediating the interaction between particles. Clearly Δx^3 represents the black body volume.

Recall now that, in agreement with the arbitrariness of n, the ranges of eqs (1,1) can be regarded as arbitrary functions of time through Δt; read for instance $\Delta x = x - x_o$ with $x = x(\Delta t)$ and $x_o = x_o(\Delta t)$, being in general $x(\Delta t)$ and $x_o(\Delta t)$ different time functions. Of course no hypothesis is necessary about these functions, which are undefined and undefinable. Hence the size of Δx is in general an arbitrary function of time itself, whereas the concept of derivative relies in the frame of eqs (1,1) only as mere ratio of uncertainty ranges. This idea generalizes the previous definition of force field $\Delta F_x = F_x^{conf} - F_x^{\infty}$. For instance $\Delta p_x/\Delta t$ takes the physical meaning of force field component $\Delta\varepsilon/\Delta x$ generated within Δx by the change rate of all p_x compatible with Δp_x during Δt, whatever the physical reason affecting p_x might be. Moreover, being the range sizes arbitrary, these ratios can even take the local physical meaning elucidated by the familiar notations $\Delta\varepsilon \to d\varepsilon$, $\Delta t \to dt$ and $\Delta p_x \to dp_x$. In other words, the local concept of derivative is here a particular case of that of ratio of arbitrarily sized uncertainty ranges. There is no contradiction between $\Delta\varepsilon/\Delta t$ and $d\varepsilon/dt$, which have both mere conceptual meaning and in fact are both indeterminable: the former because of the arbitrariness of the range boundaries, the latter because the local variables p_x and t around which shrink the respective ranges are arbitrary as well. The consistency of this position with the concept of covariancy has been concerned in [7]; in this paper and in [4] has been also shown that just the evanescent concept of distance required by the agnostic positions (1,2) in fact determines the non-locality of the quantum world. Exploit now eqs (1,1) to calculate in any reference system R an arbitrary size change $d\Delta p_x$ of $\Delta p_x = p_x - p_{ox}$ as a function of that, $d\Delta t$, of the time uncertainty range Δt, assuming that n remains constant during $d\Delta t$; hence during $d\Delta t$ the size of Δx necessarily changes by an amount $d\Delta x$ as well. Of course this reasoning can be reversed: a force field arises within the space-time range Δx because of its deformation $d\Delta x$ that in turn, because of eqs (1,1), requires the momentum range Δp_x deformation as well [7]. Is evident the link of these ideas with the foundations of relativity. Differentiating eqs (1,1) and dividing by $d\Delta t$, one finds $d\Delta p_x/d\Delta t = -(n_x\hbar/\Delta x^2)(d\Delta x/d\Delta t)$. Of course, in R' one would obtain $d\Delta p'_x/d\Delta t' = -(n'_x\hbar/\Delta x'^2)(d\Delta x'/d\Delta t')$; yet any consideration carried out about the unprimed equation can be identically carried out on the primed equation. In the present model there is no local value defined in R that changes into a new value in R', while any uncertainty range undefined in R remains undefined in R' too; so considering primed and unprimed range sizes means actually renaming a unique undefined range. The same holds of course for the ratios of any two ranges. If in particular $\Delta t = t - t_o$ is defined with constant t_o, since actually even this latter could be itself a function of t without changing anything so far introduced, then one finds in any R

$$\frac{d\Delta p_x}{dt} = -\frac{n_x\hbar}{\Delta x^2}v'_x = F_x - F_{ox}, \qquad (2,7)$$

$$F_x = \dot{p}_x, \qquad F_{ox} = \dot{p}_{ox}, \qquad v'_x = \frac{d\Delta x}{d\Delta t}.$$

Having replaced any local distance x with the uncertainty range Δx including it, the local force F_x is replaced by a corresponding range ΔF_x including local values of force. The notation n_x emphasizes that the arbitrary number n of states refers here to the x components of $\Delta\mathbf{p}$, $\mathbf{v}'$, $\mathbf{F}$ and $\mathbf{F}_o$; of course are likewise definable n_y and n_z too. Moreover note that v'_x is conceptually different from v_x introduced in section 1: despite both have formally physical dimensions of velocity, the latter only is the actual average velocity of any real particle traveling through its delocalization range Δx during Δt, the former is the deformation extent $d\Delta x$ of Δx during the time increment $d\Delta t$. So v_x is self-defined without need of further considerations, the physical meaning of v'_x is instead strictly related to that of F_x concurrently inferred. This distinction is inherent the character of the present theoretical model that, as previously remarked, concerns the uncertainty ranges of the phase space where any particle could be found rather than the particle itself; however the examples of the angular momentum and hydrogenlike energy levels have shown that working on the uncertainty ranges that define a physical property allows to gain information on the related behavior of the particle and on the given law itself. Eqs (2,7), reported here for clarity, have been early introduced in [7] and therein exploited to infer as a corollary in the particular case of constant p_{ox} (i) the equivalence principle of general relativity, (ii) the coincidence of gravitational and inertial mass and then (iii) the Newton gravity law as a particular case; actually this law results to be the first order approximation of a more general equation allowing to calculate some interesting results of general relativity, for instance the perihelion precession of planets.

Also in the present model, therefore, the deformation of the space time quantum delocalization range entails the arising of a force as a corollary of eqs (1,1). In this paper we propose a further way of handling eq (2,7): in agreement with the purpose of this paper, i.e. to infer various forms of interaction between particles from a common principle, it is enough to rewrite eqs (2,7) in different ways and examine the respective consequences. The fine structure constant α enables $\hbar$ to be eliminated from eqs (2,7), which read in c.g.s. units for simplicity

$$F_x - F_{ox} = \pm\frac{e'e}{\Delta x^2}, \qquad e' = \pm\frac{n_x v'_x}{\alpha c}e. \qquad (2,8)$$

Here $\Delta F_x = F_x - F_{ox}$ is the force field between two charges e and e' interacting through their linear charge densities $e/\Delta x$ and $e'/\Delta x$: i.e. even the electric interaction force relies on a physical basis similar to that of the gravity force. The double sign accounts for both chances that Δx expands or shrinks at deformation rate $\pm v'_x$, which is a decisive parameter to express the respective states of charge. If $v'_x = 0$ then $e' = 0$, i.e. it corresponds to a chargeless particle; of course the related electric force is null, i.e. $F_x = F_{ox}$ accounts for

other forces possibly acting on the particle, for instance the gravity; this case, concerned in [7] to emphasize the link between quantum theory and relativity, is skipped here. Moreover holds an obvious boundary condition on n_x, i.e. a value of n_x must necessarily exist such that $e' = \pm e$. Be n' this value such that by definition $n'v'_x = \alpha c$; being n'_x arbitrary integer and v'_x arbitrary as well, this position is certainly possible. Then

$$e' = \pm(n_x/n')e. \qquad (2,9)$$

Here the double sign agrees with the chances allowed for e depending on the expansion or contraction of Δx. It is reasonable to assume that $n' = 3$; considering also the deformation rates $\pm v'_y$ and $\pm v'_z$ of Δy and Δz defined likewise to v'_x, the number of states is actually counted as $n' = n_x + n_y + n_z$ with ground values $n_x = n_y = n_z = 1$, while being $1 \leq n_{xi} \leq n'$ depending on the number of respective force components $F_{xi} - F_{oxi}$ actively contributing to n'. Consider first the x-component, eq (2,7), only. If $n_x = n' = 3$, then $e'_{(3)} = \pm e$ corresponds to electron and proton charges; $F_x - F_{ox}$ of eq (2,8) is the related Coulomb force component. The case $n_x = 2$ yields $e'_{(2)} = \pm(2/3)e$, whereas $n_x = 1$ yields $e'_{(1)} = \pm(1/3)e$; accordingly $F_x - F_{ox}$ must have a characteristic physical meaning that will be concerned in section 5. The same result would be obtained considering the y or z components corresponding to eq (2,7). Hence fractional charges are in principle to be expected in nature. It is easy guess how many particles with fractional charges, the well known quarks, are to be expected. Consider the four chances corresponding to the double signs of $e'_{(1)}$ and $e'_{(2)}$ and the three deformation rates v'_x, v'_y and v'_z; the previous discussion has exemplified the link of e' with v'_x only, yet an analogous reasoning holds of course also for v'_y and v'_z. Instead three different situations are in general compatible with $e'_{(1)}$ and $e'_{(2)}$ when (i) $v'_x \neq 0$ only, (ii) $v'_x \neq 0$ and $v'_y \neq 0$ only, (iii) $v'_x \neq 0$ and $v'_y \neq 0$ along with $v'_z \neq 0$ too. Since n_x, n_y, n_z are independent and arbitrary, one could replace the second eq (2,8) for instance with $\pm n_x v'_x/\alpha c \pm n_y v'_y/\alpha c$, obtaining thus $\pm(n_x \pm n_y)/n'$ as done to infer eq (2,9); then one could combine n_x and n_y in order to obtain again ratios having the same values $\pm 1/3$ and $\pm 2/3$ previously found, but involving now both v'_x and v'_y instead of v'_x only. Analogous considerations hold for the case (iii) that involves also v'_z. In (i) the vector $\mathbf{F} - \mathbf{F}_o$ is oriented along one of the axes, here the x-axis, in (ii) it lies on one coordinate plane, here the $x-y$ plane; the components of $\mathbf{F} - \mathbf{F}_o$ arbitrarily oriented correspond in general to (iii), whereas a null vector is instead related to $\mathbf{v}' = 0$ i.e. $e' = 0$. Anyway, whatever the linear combination of v'_x, v'_y and v'_z might be, it is reasonable to think that these ways of inferring $e'_{(1)}$ and $e'_{(2)}$ are physically different from that involving v'_x only; otherwise stated, to the various ways of finding a given kind of charge correspond different particles. With the aforesaid 3 chances for each sign of $e'_{(1)}$ and $e'_{(2)}$ we expect therefore a variety of 12 particles in total. Since this number is reasonably expected to include particles and antiparticles, a sensible conclusion is that we should have 6 quarks and 6 antiquarks: for instance, to the $(n_x - n_y)e/n'$ quark charge corresponds the $(n_y - n_x)e/n'$ antiquark charge. Now the first problem is how to sort the charge signs between particles and antiparticles; in principle one could think the former as the ones having $e'_{(1)} = +e/3$ and $e'_{(2)} = +2e/3$, the latter as the ones with both negative signs. In this way, however, considering all values of charges compatible with n from 1 to n', one should conclude that in nature the mere charge signs discriminate particles and antiparticles. Since this is not the case, it is more sensible to expect that $e'_{(1)} = -e/3$ and $e'_{(2)} = +2e/3$, for instance, identify quarks whereas the inverted signs identify the corresponding antiquarks: likewise exist as a particular case particles with either integer charge whose antiparticles have either opposite charge.

Moreover if two charge states $-e/3$ and $+2e/3$ are consistent with six particles physically distinguishable, then each quark requires three chances of a new property, which is indeed well known and usually called color charge: each quark can exist in three quantum states, i.e. it can take three different color states. Being the quarks characterized by several quantum numbers, this way of justifying their number does not mean a specific color uniquely assigned to each one of them; rather it means introducing a number of internal freedom degrees of color that make two fractional charges consistent with six distinguishable particles. Anyway, since also anti-quarks exist for which hold the same considerations, three anti-colors must exist too.

Eventually, let us calculate how many kinds of bosons are necessary to describe the interactions between quarks via boson exchanges able to modify their initial color states. Consider for instance a charmed meson identically symbolized as $\{c\bar{c}\}$ or $\{\bar{c}c\}$ and assume that each boson mediating the quark interaction is specifically entrusted with changing one couple color-anticolor only: let for instance the exchange of one boson turn r into $\bar{r}$ and vice-versa. The mesons $\{c\bar{c}\}$ and $\{\bar{c}c\}$, formally obtained by quark-antiquark and antiquark-quark exchanges, are clearly identical and indistinguishable. Imagine therefore of turning all colors of c, whatever they might be, into the corresponding anticolors of $\bar{c}$, whose anticolors are at once turned into the respective colors. How many exchanges of color states into the respective anticolor states are consistent with the identity of $c\bar{c}$ and $\bar{c}c$? Given two objects, c and $\bar{c}$, each one of which can be found in three quantum states, the three colors, the trivial answer is 2^3; eight exchanges are not only enough to turn all color states of c into the respective anticolor states, which means by definition obtaining $\bar{c}$ from c, but also purposely necessary, as each single exchange generates a new quantum configuration of states physically distinguishable from that previously existing. Since a total of eight color-anticolor exchanges are required to account for as many different configurations, eight is also the number of different bosons required to make the aforesaid couple of identical mesons effectively indistinguishable. These different chances

of interaction, each one characterized by its own specific energy, should be someway correlated to and described by the existence of as many such particles representing the possible exchanges, i.e. just eight vector bosons. Also these particles are well known and usually called gluons. Is this reasoning extensible also to three-quark particles like neutron or proton? The quark-gluon plasma of these latter is necessarily more complex than that of the mesons, so the question arises whether the 8 gluons previously introduced are enough to describe also such three quark systems. Consider the proton uud and the antiproton $\overline{uu}\overline{d}$. The conversion $uu \to \overline{uu}$ has been already described. As concerns $d \to \bar{d}$, still holds an analogous reasoning: a specific kind of gluon undertakes to change one color into the anticolor, another kind of gluon does the same with another color and so on. However the kind of gluon exchanges that turns red into antired of the quark u cannot differ from that acting similarly on the quark d: it would mean that each gluon "recognizes" its own quark on which to act, i.e. we should admit that different $\delta(n/m)$ require different $\delta(n/V)$ depending on the respective m. But nothing in the previous eq (2,1) allows this conclusion, rather it seems true exactly the contrary because Δx defining V has nothing to do with m therein delocalized: indeed, as above stated, the indistinguishability of identical particles is just due to the possibility that any particle could be found in a given range. So it is more reasonable to think that each kind of gluon exchange affects a specific color, not the color of specific quark only; otherwise stated, the total number of gluons in a nucleon is greater than that in a meson without necessarily compelling a new kind of gluons, i.e. any gluon in the tree-quark system turns one specific color regardless of whether that color is of a quark d or u. This way of thinking allows that the gluons transmit the interaction between different quarks modifying their $\delta(n/m)$, i.e. their color quantum states, regardless of m. So, when counting the number of different gluons that allow the three-quark particle/ antiparticle exchanges the result is the same as that previously computed.

These short remarks are enough for the purposes of the present paper; further considerations on other properties like strangeness, isospin and so on, whose conservation rules are necessary for instance to describe the decay of complex particles consisting of two or three quarks, are well known and thus omitted here for brevity. The remainder of the paper aims to describe the fundamental interactions by implementing the ideas hitherto exposed.

3 The quantum interactions

Divide all sides of eqs (1,1) by $e^2\Delta x$ and recall that in general $\Delta p_x = (v_x/c^2)\Delta\varepsilon$. An intuitive hint to this equation, already concerned in [7] and important also for the present purposes, is quickly reported here for completeness. Let in an arbitrary reference system R a photon travel at speed c through an arbitrary delocalization range $\Delta x^{(c)}$, so that eqs (1,1) read $\Delta x^{(c)}\Delta p_x^{(c)} = n^{(c)}\hbar = \Delta t^{(c)}\Delta\varepsilon^{(c)}$; the superscripts emphasize that the ranges are sized in order to fulfill this delocalization condition during an appropriate time range $\Delta t^{(c)}$. Then $c\Delta p_x^{(c)} = \Delta\varepsilon^{(c)}$. To find how scale the sizes of the momentum and energy ranges with respect to $\Delta p_x^{(c)}$ and $\Delta\varepsilon^{(c)}$ in the case of a massive particle traveling at slower rate $v_x < c$ through $\Delta x^{(c)}$, write $\Delta x^{(c)}\Delta p_x^{(v)} = n^{(v)}\hbar = \Delta t^{(c)}\Delta\varepsilon^{(v)}$. Since neither v_x nor c appear explicitly in this equation, it is also possible to write $n^{(v)}\hbar = \Delta t^{(c)}\Delta\varepsilon^{(c)} = \Delta t^{(v)}\Delta\varepsilon^{(v)}$; this is indeed true if $\Delta t^{(c)}$ and $\Delta\varepsilon^{(c)}$ scale like $\Delta t^{(v)} = (c/v_x)\Delta t^{(c)}$, as it is reasonable, and $\Delta\varepsilon^{(v)} = (v_x/c)\Delta\varepsilon^{(c)}$. Replacing these positions in the former equation yields $\Delta x^{(c)}\Delta p_x^{(v)} = \Delta t^{(c)}(v_x/c)\Delta\varepsilon^{(c)}$ whence $c\Delta p_x^{(v)} = (v_x/c)\Delta\varepsilon^{(c)}$. Actually the superscripts can be omitted because they do not identify particular range sizes; both $\Delta p_x^{(v)}$ and $\Delta\varepsilon^{(c)}$ are indeed arbitrary like v_x itself. The superscripts are also irrelevant as concerns the functional relationship between the local values of the respective variables, which reads $p_x = (v_x/c^2)\varepsilon$ regardless of how the respective uncertainty ranges are defined. Note that p_x and ε, exactly determined in classical physics and in relativity, are instead here random values within the respective uncertainty ranges. Also note that an identical reasoning in R' solidal with the particle would yield $p'_x = (v'_x/c^2)\varepsilon'$: this is therefore a quantum expression relativistically invariant. This kind of reasoning has been carried out in [7] to show the connection between quantum mechanics and relativity. Now instead consider for the next discussion the following equations directly inferred from eqs (1,1)

$$\frac{n\hbar v_x}{\Delta x} = \Delta\varepsilon, \qquad v_x = \frac{\Delta x}{\Delta t}, \qquad v_x \le c. \tag{3,1}$$

The last position does not merely emphasize a feature in principle expected for any velocity, it takes a special relevance in the present context. Being $\Delta\varepsilon$ and Δx arbitrary, one could write $\Delta p_x = \Delta\varepsilon^o v_x^o/c^2$ too, with v_x^o and $\Delta\varepsilon^o$ still fulfilling the given Δp_x. The total arbitrariness of the range sizes plays a key role in the following reasoning based on $v_x\Delta\varepsilon = v_x^o\Delta\varepsilon^o$: if $v_x = c$, then necessarily $v_x^o < c$ and $\Delta\varepsilon^o > \Delta\varepsilon$. Examine step by step this point writing identically eq (3,1) as follows

$$\frac{e^2}{\Delta x} = \frac{\alpha}{n}\frac{v_x^o}{c}\Delta\varepsilon^o, \qquad \frac{v_x v_x^o}{c^2} = \frac{\Delta\varepsilon}{\Delta\varepsilon^o}, \qquad \Delta\varepsilon \le \Delta\varepsilon^o. \tag{3,2}$$

The last position emphasizes that both chances $\Delta\varepsilon^o = \Delta\varepsilon$ and $\Delta\varepsilon^o \neq \Delta\varepsilon$ are equally possible. If $\Delta\varepsilon = \Delta\varepsilon^o$, then $v_x = v_x^o$ compels concluding $v_x = v_x^o = c$ only; so eqs (2,7) and (3,2) yield $e^2/\Delta x = \chi\Delta\varepsilon$, being $\chi = \alpha/n$ a proportionality factor. This means correlating the potential energy $e^2/\Delta x$ of two electric charges to $\Delta\varepsilon$, introduced through Δp_x and thus having the meaning of kinetic energy range. On the one hand $\Delta\varepsilon^o \neq \Delta\varepsilon$ requires different v_x^o and v_x, thus both velocities or at least either of them smaller than c, whence the inequality; on the other hand, relating the physical meaning of the velocities hitherto introduced to that of the boson carriers that

mediate the interaction force between particles: $v_x^o = c$ requires massless bosons, $v_x^o < c$ massive bosons. Therefore the arbitrariness of $\Delta\varepsilon$ and $\Delta\varepsilon^o$ justifies the conclusion that either chance of range sizes prospects different results for eqs (3,2) and (3,1), despite their common origin from eqs (1,1). Two questions arise at this point: (i) whether these equations describe two different interactions or two different appearances of a unique interaction, (ii) whether or not it is possible to infer from both equations a relationship like $e^2/\Delta x = \chi\Delta\varepsilon$ despite their formal difference. The answers rely on the fact that in eq (3,2) appears explicitly the Coulomb charge e inherent the definition of α, in eq (3,1) it does not necessarily hold; nothing compels assuming that even the energy $n\hbar v_x/\Delta x$ is by necessity referable to a Coulomb energy.

If $n\hbar v_x/\Delta x$ does, then the common origin of these equations from eqs (1,1) is a good reason to expect that the chances of massive or massless vector bosons are merely two different ways of manifesting a unique kind of interaction; rewriting the inequality as $\Delta\varepsilon^o = \Delta\varepsilon + \delta\varepsilon$, with $\delta\varepsilon \geq 0$ of course arbitrary like $\Delta\varepsilon$ and $\Delta\varepsilon^o$, both chances are in principle acceptable depending on the amount of energy at which the interaction occurs. In other words $\delta\varepsilon > 0$ is an additional energy range motivated by the arbitrariness of $\Delta\varepsilon$, which indeed admits introducing also $\Delta\varepsilon^o$ too, and justifying the presence of massive vector bosons. By consequence the chance of finding a unique link like $e^2/\Delta x = \chi\Delta\varepsilon$ between potential and kinetic energies is to be reasonably expected; so, fixing an arbitrary $\Delta\varepsilon$ allows assessing via χ the relative strengths of both interactions at comparable values of $\Delta\varepsilon$ and respective characteristic lengths Δx. The physical consequences of this reasoning are exposed in section 4.

If instead $n\hbar v_x/\Delta x$ is an energy not referable to that between integer charges, in fact nothing hinders thinking that it is directly related to the aforesaid fractional charges; according to eq (2,8), $v_x = \Delta x/\Delta t$ is physically different from $v'_x = d\Delta x/d\Delta t$. Then eq (3,1) describes an interaction prospectively different from that of eq (3,2); so the former equation must be considered regardless of the latter to check what kind of physical information follows from the considerations of section 2. Also the consequences inferred from these equations are expectedly different; in particular the link χ between potential and kinetic energies should be reasonably different in either case just mentioned. In other words, χ can be compared for similar $e^2/\Delta x$ and $\Delta\varepsilon$ to characterize the relative strengths of the various kinds of interactions. The physical consequences of this reasoning are exposed in section 5.

These are the key ideas to be further highlighted below. The dual way of elaborating a unique principle, the statistical formulation of quantum uncertainty, has an intrinsic physical meaning coherent with the purposes of the present paper, i.e. to demonstrate that kinds of interaction apparently different are in fact consequences of a unique principle. In other words, eqs (3,2) and (3,1) are the starting point to distinguish two cases, which will be discussed separately under the only conceptual constraint of being mutually self-consistent. The following sections 4 and 5 aim to outline the respective ways to link the potential and kinetic energies.

4 The interaction according to eqs (3,1) and (3,2)

The following discussion concerns the ways to reduce the eqs (3,1) and (3,2), regarded together, to the form $e^2/\Delta x = \chi\Delta\varepsilon$ in both cases $\delta\varepsilon = 0$ and $\delta\varepsilon > 0$. Consider first $\delta\varepsilon = 0$, which requires $v_x^o = v_x = c$ and thus massless boson carriers. So the unique result possible is

$$\frac{e^2}{\Delta x} = \chi_{em}\Delta\varepsilon, \qquad \chi_{em} = \frac{\alpha}{n}. \tag{4,1}$$

Here α/n emphasizes the electromagnetic interaction in analogy with eq (1,4).

The further chance $\delta\varepsilon > 0$ requiring the condition $v_x^o < c$ prospects instead the presence of massive boson carriers; thus $\delta\varepsilon > 0$, related to the formation of massive carriers, represents reasonably the energy gap with respect to the former case of eq (4,1) involving massless carriers only. While heavy vector bosons are the physical consequence of the concurring inequalities $v_x^o < c$ and $\delta\varepsilon > 0$, the arbitrariness of v_x^o prevents the possibility of deciding a priori either chance for $\delta\varepsilon$ and compels the conclusion that a unique kind of interaction is actually compatible with both chances. It will be shown that the interaction energy related to the possible size of Δx discriminates either chance. Despite both chances are incorporated into a unique conceptual frame, further considerations are necessary in this case. Write the first eq (3,2) as follows

$$\frac{e^2}{\Delta x} = \frac{\alpha^2}{n^2}\frac{\Delta\varepsilon^o}{q^o}, \qquad q^o = \frac{e^2}{n\hbar v_x^o}, \qquad v_x < c. \tag{4,2}$$

Since eqs (3,2) require $\Delta\varepsilon^o/q^o = (c/v_x)(n/\alpha)\Delta\varepsilon$, the obvious inequality

$$(n/\alpha)^2 > v_x/c \tag{4,3}$$

yields $\Delta\varepsilon^o/q^o > (\alpha/n)\Delta\varepsilon$. Hence a value $q_w > q^o$ certainly exists such that

$$\Delta\varepsilon^o/q_w = (\alpha/n)\Delta\varepsilon. \tag{4,4}$$

Replacing this result into the first eq (4,2), one finds

$$\frac{e^2}{\Delta x_w} = \chi_w\Delta\varepsilon, \qquad \chi_w = \left(\frac{\alpha}{n}\right)^3, \qquad \Delta x_w = \frac{q_w}{q^o}\Delta x. \tag{4,5}$$

The first equation is formally analogous to eq (4,1) a scale factor q_w/q^o for Δx apart, while α/n is replaced by the much smaller quantity $(\alpha/n)^3$; hold however for χ_w considerations analogous to that previously carried out for χ_{em}, i.e. it links kinetics and potential energies. The explicit form of the inequality (4,3) reads $(n\hbar c)^2 > e^4(v_x/c)$, so that $(n\hbar c/\Delta x)^2 > (e^2/\Delta x)^2(v_x/c)$ and thus $(c\Delta p_x)^2 > (e^2/\Delta x)^2(c\Delta\varepsilon/v_x^o\Delta\varepsilon^o)$; as $c\Delta p_x = \Delta\varepsilon^o v_x^o/c$, i.e. $c\Delta p_x = (q_w v_x^o\alpha/nc)\Delta\varepsilon$ according to eq (4,4), the inequality (4,3) reads

$$(\zeta\Delta\varepsilon)^3 > (e^2/\Delta x)^2\Delta\varepsilon, \qquad \zeta = \frac{q_w v_x^o\alpha}{nc} = w\left(\frac{n}{\alpha}\right)^2. \tag{4,6}$$

Hence an energy $\varepsilon_0 > 0$ certainly exists such that

$$\zeta^3\Delta\varepsilon^3 - (e^2/\Delta x)^2\Delta\varepsilon - \varepsilon_0^3 = 0. \quad (4,7)$$

Regarding ζ as a constant through an appropriate choice of q_w, not yet specified and here accordingly defined, let us solve the eq (4,7) in order to introduce three real sizes $\Delta\varepsilon_j$, $j = 1, 2, 3$. Note that this does not mean assigning definite values to the size of $\Delta\varepsilon$, which remains indeed arbitrary and unknown like any uncertainty range because of Δx; solving eq (4,7) means examining the physical information consistent with some particular range sizes that fulfil the inequality (4,6). One finds

$$\varepsilon_0 = \left(\frac{2\sqrt{3}}{9}\right)^{1/3}\zeta^{-1/2}\frac{e^2}{\Delta x}, \qquad \Delta\varepsilon_1 = \frac{2}{\zeta^{3/2}\sqrt{3}}\frac{e^2}{\Delta x}, \quad (4,8)$$

$$\Delta\varepsilon_{2,3} = \Delta\varepsilon_2 = \Delta\varepsilon_3 = -\frac{1}{\zeta^{3/2}\sqrt{3}}\frac{e^2}{\Delta x}.$$

The former equation is the condition to make null the imaginary parts of the roots $\Delta\varepsilon_2$ and $\Delta\varepsilon_3$ that, as emphasized by the last equation, result by consequence coincident. As expected, all quantities expressed here as a function of Δx are in fact arbitrary like this latter. The constant ζ can be eliminated from the equations; so

$$\frac{\Delta\varepsilon_1}{\varepsilon_0} = \frac{2}{\zeta\sqrt{3}}\left(\frac{9}{2\sqrt{3}}\right)^{1/3}, \qquad \frac{\Delta\varepsilon_1}{\varepsilon_0^3} = \frac{3}{(e^2/\Delta x)^2}, \quad (4,9)$$

$$\frac{e^2}{\Delta x} = \varepsilon_0\sqrt{3\frac{\varepsilon_0}{\Delta\varepsilon_1}}.$$

It is interesting to rewrite eq (4,7) as $(\zeta^3\Delta\varepsilon^2 - (e^2/\Delta x)^2)\Delta\varepsilon = \varepsilon_0^3$, which yields

$$\Delta t = n\hbar\zeta^3\frac{(e^2/\Delta x)^2}{\varepsilon_0^3}\left[\left(\frac{\Delta\varepsilon}{e^2/\Delta x}\right)^2 - \zeta^{-3}\right].$$

In this way $\Delta\varepsilon^3$ splits into a multiplicative factor $\Delta\varepsilon$, related to Δt through eqs (1,1), times a factor merging together $\Delta\varepsilon^2$ and $(e^2/\Delta x)^2$. Let us specify in particular Δx as Δx_w of eq (4,5); owing to the last eq (4,9), one finds then

$$\Delta t_w = \frac{3n\hbar\zeta^3}{\Delta\varepsilon_{1w}}\left[\left(\frac{n^3}{\alpha^3}\right)^2 - \zeta^{-3}\right], \quad \Delta\varepsilon_{1w} = \frac{2}{\zeta^{3/2}\sqrt{3}}\frac{e^2}{\Delta x_w}. \quad (4,10)$$

Despite Δx is unknown and arbitrary by definition, when it is specified as the range Δx_w purposely pertinent to eq (4,5) the former equation takes the form $\Delta t \propto (n/\alpha)^6$ plus a term $\tau = 3n\hbar/\Delta\varepsilon_{1w}$. If $\Delta\varepsilon_{1w}$ and n are large enough so that $\tau << (n/\alpha)^6$, then Δt_w and the factor χ_w linking $e^2/\Delta x_w$ and $\Delta\varepsilon$ of eq (4,5) fulfill the well known condition

$$\Delta t_w \propto \chi_w^{-2}.$$

Note now that

$$\Delta\varepsilon_1 + \Delta\varepsilon_2 + \Delta\varepsilon_3 = 0 \quad (4,11)$$

and that eq (4,7) is directly related to $v_x^o/c < 1$ because it comes from the inequalities (4,2) and (4,3). Moreover each energy range by definition introduces its own random value of energy; this suggests that are related to eq (4,5) three characteristic energies, i.e. three corresponding massive particles, whose energies are by definition included within the uncertainty ranges of eqs (4,11).

Consider in general three energy ranges $\Delta\varepsilon_j = \varepsilon_j' - \varepsilon_j''$, being $j = 1..3$, of course with both ε_j' and ε_j'' arbitrary and unknown; define then the energies η_j included within them as $\eta_j = (\varepsilon_j' + \varepsilon_j'')/2$, i.e. as average values of the respective boundary values. It is immediate to realize that the condition $\sum\Delta\varepsilon_j = 0$ is compatible with $\sum\eta_j \neq 0$; indeed $\sum(\varepsilon_j' - \varepsilon_j'')/2 = 0$ reads identically $\sum(\varepsilon_j' + \varepsilon_j'')/2 - \sum\varepsilon_j'' = 0$, whence in general $\sum\eta_j = \sum\varepsilon_j'' \neq 0$. Repeat this reasoning regarding η_j as the average values of the specific energy ranges of eq (4,11). The fact that $\eta_{tot} = \eta_1 + \eta_2 + \eta_3 \neq 0$ agrees with the idea of interaction energy; indeed no constrain could be definable for three independent free particles. On the one hand the chance of replacing any quantum range with its average, as done here for $\Delta\varepsilon_j$ and η_j, has a general valence because the range sizes are arbitrary, undefined and undefinable like the average value inferred from their boundaries. Since any value allowed to the former is also allowed to the latter, considering η_j instead of $\Delta\varepsilon_j$ does not exclude the point of view of eqs (1,1): replacing an arbitrary value with another arbitrary value corresponds to replace n with n', which is however immaterial because both symbolize sets of integer values and not specific values. On the other hand the ranges (4,11), regarded all together, fulfill globally the energy conservation regardless of whether $\Delta\varepsilon_j \neq 0$ or $\Delta\varepsilon_j = 0$; as just shown, however, the same does not necessarily hold for η_{tot}. To make also this latter compliant with the eq (4,11), let us assume therefore that η_{tot} has a finite lifetime of the order of $\hbar/\eta_{tot}$. Let Δt_w be this lifetime. In agreement with eq (4,10), during Δt_w the sum $\sum\Delta\varepsilon_j$ is still globally null likewise as before and after their actual transient appearance; in this way the massive particles concerned by the respective energy ranges are jointly involved as concurrent physical properties inherent eq (4,5) and thus the present kind of interaction. The physics of the weak interactions is well known. Here, as a significant check of these ideas, we propose a simple energy balance to infer the energies η_j and thus η_{tot} exploiting just the requirement that the η_j must be regarded all together.

A possible interpretation of the equal sizes and negative signs of $\Delta\varepsilon_2$ and $\Delta\varepsilon_3$, despite in the present model the ranges are always introduced positive by definition, is that their sum with $\Delta\varepsilon_1$ equal to zero requires interacting particles; as explained in section 2, no relationship would be possible by definition for free particles. Let two of them, say η_2 and η_3,

interact in order to release the energy necessary to form also η_1. The fact that $\eta_2 = \eta_3$ because of $\Delta\varepsilon_2 = \Delta\varepsilon_3$ means that their interaction occurs regarding identically either of them in the field of the other one; together, therefore, these particles provide the energy necessary to allow the kind of interaction here concerned. The simplest hypothesis is that the particles η_2 and η_3 have charges of opposite signs whereas η_1 is neutral, thus fulfilling the global charge conservation before, during and after their lifetime; if so, the energy gain of Coulomb energy at an appropriate interaction distance justifies also the neutral particle η_1. In this way the model allows the existence of three range sizes whose finite lifetimes agree with the finite values of the respective heavy bosons η_j. This conclusion is summarized as follows

$$\eta_{tot} = \eta_1 + (\eta_2 + \eta_3), \qquad \eta_2 = \eta_3.$$

The second equation emphasizes that actually η_2 and η_3 form a Coulomb system of charges, whose energy transient uniquely defined likewise ε_{el} of eq (1,4) characterizes the present kind of interaction. This idea suggests to estimate η_{tot} just computing the energy levels of the system of charges η_2 and η_3 by analogy with that of a hydrogenlike atom. Exploit for simplicity the previous non-relativistic equations (1,3) and (1,4); owing to the generality of these equations, there is no reason to exclude that analogous considerations hold at least approximately also here putting of course the charge $Z = 1$ and describing the system of charged bosons η_2 and η_3 as due to $-\varepsilon_{el} = \pi e^2/n\lambda$. It is necessary to take into account however that now also the neutral particle η_1 contributes to η_{tot}, in agreement with the idea of regarding the particles all together. Guess first according to eqs (4,8) that the mass of η_1 should have the same order of magnitude of η_2 and η_3, so that $\eta_{tot} \approx 3\eta_2$; the chance of identifying η_{tot} with $-\varepsilon_{el}$ is consistent with this idea simply putting

$$\eta_{tot} = \pi\eta_2, \quad \eta_2 = \eta_3 = e^2/n\lambda, \quad \eta_1 = (\pi - 2)e^2/n\lambda. \quad (4,12)$$

In other words, eq (1,4) suggests that the expected coefficient ≈ 3 must be actually regarded as π. Despite the non-relativistic reasoning, these conclusions are correct because confirmed by the experience. The experimental masses of the $W^{\pm}$ and Z^0 vector bosons are $m_{W^{\pm}} = 80.39$ GeV and $m_{Z^0} = 91.19$ GeV respectively, for a total mass of $m_{tot} = 251.97$ GeV; in effect

$$m_{tot} = 3.134\, m_{W^{\pm}} \qquad m_{Z^0} = 1.134\, m_{W^{\pm}}$$

are compatible with the values expected for π and $\pi - 2$. Trivial considerations show that the reduced Compton lengths $\bar{\lambda}$ of the vector bosons consistent with $e^2/n\lambda$ are $\bar{\lambda}_{\eta_1} = n\lambda/((\pi - 2)\alpha)$ and $\bar{\lambda}_{\eta_2} = \bar{\lambda}_{\eta_3} = n\lambda/\alpha$, having introduced explicitly the masses $m_j = \eta_j/c^2$. These results are confirmed considering the zero point energy $\Delta p_j^2/2m_j$ of the vector bosons η_j, where $\Delta p_j = p_2 - p_1$ is the gap between its momentum p_2 after confinement within a given delocalization range Δx_w and its initial momentum p_1 in an ideal unconfined state; hence the corresponding energy gap after confinement within Δx_w resulting from the x, y and z components is $\Delta p_j^2/2m_j = 3(n^2\hbar^2/2m_j\Delta x_w^2)$. Assume now that the confinement energy $\Delta p_j^2/2m_j$ is just the energy $\eta_j = m_j c^2$ itself that determines the space-time scale of this kind of interaction, i.e.

$$\eta_j = \frac{3}{2}\frac{c^2\Delta p_j^2}{\eta_j}; \qquad (4,13)$$

then $\Delta x_w = (3/2)^{1/2} n\hbar c/\eta_j$, i.e. for η_2 and η_3

$$\Delta x_w = (3/2)^{1/2} n^2 \lambda(\hbar c/e^2). \qquad (4,14)$$

For $n = 1$ therefore Δx_w coincides with $\bar{\lambda}_{\eta_2} = \bar{\lambda}_{\eta_3}$ a trivial numerical factor $\sqrt{3/2}$ apart; an identical conclusion holds of course for η_1 too, the numerical factor $(\pi - 2)\sqrt{3/2}$ apart. This confirms the assumed link between delocalization extent and energy of the force carriers, which allows identifying $\eta_{tot} = -\varepsilon_{el}$ in agreement with eq (1,3).

Put first $n = 1$ in eqs (4,12). The value of λ corresponding to the energies of the particles η_2 and η_3 is $\lambda = 1.79\times10^{-20}$ m, so that $\bar{\lambda}_{\eta_2} = \bar{\lambda}_{\eta_3} = 2.45 \times 10^{-18}$ m and $\bar{\lambda}_{\eta_1} = 2.15 \times 10^{-18}$ m; the characteristic range Δx_w of interaction is thus of the order of 10^{-18} m. Since the classical proton radius $r_p = e^2/m_p c^2$ is about 0.8768 fm according to recent measurements [13], the above energies concern a sub-nuclear scale interaction; vice-versa, one could estimate the correct scale of energy of the vector bosons requiring an interaction that occurs at the sub-nuclear extent at which one calculates $\chi_w = \alpha^3 = 3.9 \times 10^{-7}$.

So far we have considered $n = 1$. What however about $n > 1$? First of all, Δx_w becomes n times larger than the aforesaid Compton lengths of η_j; this deviation means a longer range allowed to the interaction. Moreover, according to eqs (4,12) $\eta_{tot} \to 0$ for $n \to \infty$; at this limit the aforesaid space scale of interaction is inconsistent with the corresponding energies of massive boson carriers, which therefore should expectedly require an appropriate threshold energy to be activated. For $n \to \infty$ is thus allowed the less energy expensive and longer range interaction with $\delta\varepsilon = 0$ only, in agreement with the initial idea that $\delta\varepsilon \neq 0$ is related to the boson masses. This conclusion is intuitively clear, but what about the energy threshold? According to the eqs (4,12) the energies η_1, η_2 and η_3 downscale with n, whereas according to eq (4,14) Δx_w upscales with n^2; so the lower threshold for the existence of massive bosons, i.e. for the validity of these equations themselves, concerns n of $\eta_{tot}^{(n)} = -\varepsilon_{el}(Z = 1, n) = (\pi/n)e^2/\lambda$: it is required that the interaction distance of the hydrogenlike system of charges enable the energy to create vector bosons. The inequality $\eta_{tot}^{(n)} > e^2/\lambda$, which holds for $n \leq 3$, ensures that, whatever the masses $\eta_2^{(n)}$ and $\eta_3^{(n)}$ might be, the energy gain due to their Coulomb interaction accounts not only for the energy e^2/λ of the system of charged particles themselves but also for the surplus required by the neutral particle $\eta_1^{(n)}$. Clearly the threshold corresponds to the

value $\eta^{(3)}_{tot} = (\pi/3)e^2/\lambda$, i.e. about 81 GeV; the corresponding Compton lengths of the bosons are $\lambda^{(3)}_{\eta_2} = \lambda^{(3)}_{\eta_3} = 3\lambda/\alpha$ and $\lambda^{(3)}_{\eta_1} = 3\lambda/((\pi-2)\alpha)$. In fact even for $n = 3$ these lengths are of the order of 10^{-17} m, i.e. still consistent with a subnuclear range. At energy below this threshold, i.e. $n \geq 4$, eq (4,1) only describes the interaction. Of course the most favorable condition for this interaction to occur is that with $n = 1$, which ensures the maximum binding energy given by eq (4,12) and corresponds to the shortest interaction distance and maximum values of the three boson masses in fact experimentally detected. The model admits however even the possible existence of lighter bosons. In conclusion, the different energy scales characterize the features of eqs (4,1) or (4,5) because of different values of n; both equations describe however the same kind of interaction.

5 The interaction according to eq (3,1)

The starting point of this section is the eq (3,1) that reads

$$\frac{n\hbar v_x}{\Delta x} = \chi_s \Delta\varepsilon, \qquad \chi_s = 1. \tag{5,1}$$

The lack of coefficient at right hand side of eq (3,1) is tentatively interpreted here as the presence of coefficient $\chi_s = 1$. Being v_x and n arbitrary, it is certainly possible to introduce a proportionality constant ξ defined as $n\hbar v_x = \xi e^2$; so eq (3,1) reads $(\xi e^2/\Delta x)/\Delta\varepsilon = 1$. Usually a proportionality constant linking two quantities that fulfill a given condition or a given physical law is of the order of the unity, unless some specific reason compels an appropriate hypothesis about its actual order of magnitude. Since here even Δx and $\Delta\varepsilon$ are arbitrary, however, it is difficult to guess a valid reason to compel ξ very different from the unity. So, in terms of order of magnitude, the position $\xi \approx 1$ seems reasonable although not thoroughly demonstrated, whence the tentative conclusion quoted in eq (5,1). On the other hand, once having reduced this equation to the form $(e^2/\Delta x)/\Delta\varepsilon = \chi_s$, one can compare $\chi_s = \xi^{-1} \approx 1$ with $\chi_{em} \approx \alpha$ and $\chi_w \approx \alpha^3$ defined by the equations (4,1) and (4,5) formally similar, of course under the assumption that the ranges at left hand sides defining these values are comparable as well. Even without a specific reason to exclude the plain idea $\chi_s \approx 1$, a better assessment of this conclusion appears however necessary: the lack of e^2 at left hand side, replaced by $n\hbar v_x$, allows handling eq (5,1) in order to introduce the interaction between the fractional charges concerned in section 2; but this chance, suggested by eqs (2,8) and (2,9) that anyway do no exclude themselves $\xi \approx 1$, is justified only revising the term $e^2/\Delta x$.

Consider again the eq (2,7) $F_x = -a'/\Delta x^2 + F_{ox}$ with $a' = n\hbar v'_x$ in the simplest case where both a' and $\dot{p}_{ox} = F_{ox}$ are constants. Actually these constants could likely be first order approximations only of series developments whose higher order terms are neglected; yet, even this approximate meaning of the eq (2,7) is enough for the present discussion. Assuming $F_{ox} < 0$ likewise as the first addend in order to describe an attractive force, F_x is compatible with a potential energy U_i of the i-th quark having the form

$$U_i = -\frac{a}{\Delta x} + b\Delta x + U_0 \tag{5,2}$$

being U_0, a and b appropriate integration constants; the latter is clearly related to F_{ox}. Considering $\Delta U_i = U_i - U_0$ one recognizes a well know formula, the so called "asymptotic freedom", describing the interaction between quarks; of course in the present model where any local distance x randomly included by its quantum uncertainty range is replaced by a range of distances Δx, the local value of potential energy U_i turns into a range ΔU_i of allowed values. Let us examine the eq (5,2) in two particular cases where (i) $a/\Delta x \approx b\Delta x$ and (ii) $a/\Delta x \approx U_0$; the arbitrary size of Δx justifies in principle both chances. The former case holds when $\Delta x^{(\mathrm{i})} \approx \sqrt{a/b}$ and yields $U^{(\mathrm{i})}_i \approx U_0$; according to the chance (ii) $\Delta x^{(\mathrm{ii})} \approx a/U_0$ yields instead $U^{(\mathrm{ii})}_i \approx b\Delta x^{(\mathrm{ii})} = ba/U_0$. This means that a delocalization extent of the system quark + gluons around $\Delta x^{(\mathrm{i})}$ the potential energy is approximately of the order of U_0, around a range $\Delta x^{(\mathrm{ii})}$ the potential energy increases linearly with Δx. Define a and b in agreement with eqs (2,8) and (2,9) in order that eq (5,2) takes a reasonable form. Put a proportional to the electric charge $c_i^2 = (\pm(n_i/n')e)^2$, i.e. $a = a_o c_i^2$ via the proportionality constant a_o; also, let analogously be b proportional to the color quantum number C_j, i.e. $b = b_o C_j^2$ with $j = 1 \cdot \cdot 3$. The subscripts symbolize the i-th quark in the j-th color quantum state; in this way $b = 0$ for a colorless Coulomb particle with $n_i = n'$, in which case the eq (5,2) turns, according to eq (2,9), into the classical potential energy $-e^2/\Delta x' + U_0$ of two Coulomb charges attracting each other. This reasoning suggests that the color quantum number should have the form $C_j = f_{j1}(n' - n_i)^2 + f_{j2}(n' - n_i)^4 + \cdot\cdot$, where f_{j1} and f_{j2} are appropriate coefficients of series expansion fulfilling the actual value of C_j whatever it might be; it is interesting the fact that the electric charge depends on n_i/n', the color charge on $n' - n_i$. As concerns $\Delta x' = \Delta x/a_o$, note that multiplying the size of Δx by any factor yields a new range still arbitrary and thus still compliant with eqs (1,1); for the same reasons introduced in the previous section, i.e. because any size possible for Δx is allowed to $\Delta x'$ as well, the notation $\Delta x'$ means in fact nothing else but renaming Δx. In summary, the Coulomb potential appears to be a particular case of eq (5,2), whose local features are described by the aforesaid chances; the expressions of $U^{(\mathrm{i})}_i$ and $U^{(\mathrm{ii})}_i$ are

$$\begin{gathered} \Delta x^{(\mathrm{i})} = \sqrt{\frac{a}{b}}, \qquad U^{(\mathrm{i})}_i = U_0, \\ a = l_a\varepsilon_a\left(\frac{c_i}{e}\right)^2, \qquad b = \frac{\varepsilon_b}{l_b}C_j^2, \\ \Delta x^{(\mathrm{ii})} = \frac{a}{U_0}, \\ U^{(\mathrm{ii})}_i = \frac{ab}{U_0} = \frac{\varepsilon_a\varepsilon_b}{U_0}\frac{l_a}{l_b}\left(\frac{c_iC_j}{e}\right)^2 = b\Delta x^{(\mathrm{ii})}. \end{gathered} \tag{5,3}$$

The constant energies ε_a and ε_b together with the constant lengths l_a and l_b describe the physical dimensions of a and b without need of proportionality factors. Note that $l_b \to \infty$, compels $\Delta x^{(\mathrm{i})} \to \infty$ and $b \to 0$; as the color is introduced by b, this agrees with a constant Coulomb potential $U_i^{(\mathrm{i})} = U_0$ of a colorless particle. By definition therefore $l_a\varepsilon_a = e^2$ for $n_i/n' = 1$, whereas it is expected to take a different value for $n_i/n' < 1$: the new value of $l_a\varepsilon_a/e^2$ when e^2 is replaced by $(n_i/n')e^2$ is known in the literature as $\alpha_s \approx 1$. In summary, eqs (5,3) yield

$$U_i^{(\mathrm{i})} = U_0, \quad U_i^{(\mathrm{ii})} = U_0'\left(\frac{c_i}{e}\right)^2, \quad U_0' = \frac{\alpha_s\varepsilon_b e^2 C_j^2}{l_b U_0}. \qquad (5,4)$$

Appears here once more the importance of the delocalization range Δx: in eq (4,14) Δx_w controlled either appearance of the electroweak interaction, in eqs (5,3) two different range sizes $\Delta x \approx \Delta x^{(\mathrm{i})}$ or $\Delta x \approx \Delta x^{(\mathrm{ii})}$ emphasize either feature of U_i: in (ii) it depends upon the fractional charge, in (i) it does not because $-a/\Delta x$ is balanced by $b\Delta x$ despite both terms describe attractive force.

Let us concern now eq (5,2) in a more general way. The features of U_i as a function of Δx are related to $\delta[(n/m)(n/V)]$ because Δx defines V, eq (2,1), and also because the eq (5,2) comes directly from ΔF_x of eq (2,7). What is distinctive here with respect to the gravitational or Coulomb interaction is the mere fact of having put $F_{ox} \neq 0$; so the consequent form of U_i with $b \neq 0$ describes a peculiar kind of attractive force that increases with Δx. Another remarkable point is that ΔF_x is not necessarily that between different quarks only, because eq (2,7) concerns a mere effect of confinement that holds even for an isolated quark; rather it seems more appropriate to think that the interaction between different quarks strictly replicates an intrinsic feature of the potential energy due to the confinement effect even of a single particle, which also involves its messenger bosons. In fact, in the present model Δx is by definition the delocalization range of one particle; the arising of any form of interaction is due to the presence of a further particle that possibly shares the same delocalization range. In general the number of states within a system of interacting particles is related to their energy, to their masses and to the whole space volume in which they are delocalized: eq (2,2) shows indeed that if n_1 is the number of states of the system with its particles supposed non-interacting, then δn is the change consequent to their interaction, while $\Delta\varepsilon_{n_1+\delta n}$ is the concurrent energy change from the initial $\Delta\varepsilon_{n_1}$. According to the considerations of section 2, in the present case V is the time space delocalization volume of one quark and its interaction messengers, the gluons. If a further quark could share this V, then the quarks interact. If the delocalization volume V is filled with gluons of both quarks mediating their interaction, then the change $\delta(n/V)$ stimulates a question: are the particles that mediate the interaction interacting themselves? Clearly, from the standpoint of eqs (2,7) and (5,2) this question holds even for one quark only within V. A positive answer would explain why ΔF_x increases when pulling apart the interacting quarks, e.g. of a nucleon or meson, or even a lonely quark and its gluon system; in the latter case a greater delocalization range describes indeed the chance of mowing away the gluons from their own quark, which however increases the energy of the system. To emphasize how the position $F_{ox} \neq 0$ answers the question, suppose that the quark-gluon and gluon-gluon interactions does not allow distinguishing the interaction between a quark and "its own" gluons from that of these latter with another identical quark; this would mean distinguishing identical particles, which is however forbidden by eqs (1,1) [7]. If the gluons are not mere interaction messengers but rather self-interacting messengers, then eq (5,2) describes the asymptotic freedom simply as a feature of one quark and its own system of gluons, i.e. even without necessarily requiring a further quark; otherwise stated, a net splitting of gluons from a quark interferes even with their propensity to follow another quark. The concept of asymptotic freedom is linked to the energy constrain that explains why do not exist bare quarks without gluons and bare gluons without quarks. Calculate the change of U_i as a function of Δx as $\Delta U_i = (\partial U_i/\partial\Delta x)\Delta x$ at the first order; the force field $\Delta F_x = -\partial U_i/\partial\Delta x$ acting on quark and its gluon system delocalized in Δx can be calculated in particular at the delocalization extents $\Delta x^{(\mathrm{i})}$ or $\Delta x^{(\mathrm{ii})}$. Replacing here the previous results, one finds $\Delta F_x^{(\mathrm{i})} = -2b$ and $\Delta F_x^{(\mathrm{ii})} = -b(1 + U_0/U^{(\mathrm{ii})})$.

It will be shown in the next section that $U_0' \approx 2U_0 \approx 1\,\mathrm{MeV}$; so, being U_i a monotonic function of Δx, results $\Delta x^{(\mathrm{ii})} \lesssim \Delta x^{(\mathrm{i})}$ because $U_i^{(\mathrm{ii})} \lesssim U_i^{(\mathrm{i})}$ according to eq (5,4). If $\Delta x^{(\mathrm{ii})}$ is of the order of the proton radius, i.e. 10^{-15} m, then according to eq (5,3) b results of the order of 1 GeV/fm, as it is well known. Then, inside a proton the force field at (i) is about twice than that at (ii); of course $b\Delta x$ further increases for $\Delta x > \Delta x^{(\mathrm{i})}$, i.e. outside the actual radius of the proton. This means that extending delocalization range of the quark/gluon system from $\Delta x^{(\mathrm{ii})}$ to $\Delta x^{(\mathrm{i})}$ and then to any $\Delta x > \Delta x^{(\mathrm{i})}$, i.e. allowing quark and gluons to have more space to move apart each other, corresponds to a greater energy; this is not surprising once having found that $U_i^{(\mathrm{ii})}$ is already in the region of linear increase of U_i as a function of Δx. The dependence of U_i on Δx is trivially self-evident; the reasoning about $\Delta x^{(\mathrm{ii})}$ and $\Delta x^{(\mathrm{i})}$ allows to quantify this evidence with specific reference to the sub-nuclear length scale.

The behavior of U_i and the concept of asymptotic freedom equation are straightforward consequences of eq (2,7) and thus of eqs (1,1); this feature of the strong interaction is indeed characterized by the concept of uncertainty, which in particular prevents specifying the actual size of Δx. From the present standpoint only, therefore, no kind of correlation appears in principle between quark generations and chances (i) and (ii) inherent the eq (5,2). Yet, it seems intuitive that either chance for Δx and thus either behavior of potential energy should be selectively related to the energies characteristic of

the three generations of quarks. This supposition will be confirmed in the next section, at the moment one must only admit that both chances are allowed to occur.

Now let us revert to the opening question of this section, i.e. how to regard the energy term $n\hbar v_x/\Delta x$ and χ_s of eq (5,1). The conceptual analogy of χ_s with χ_{em} and χ_w of eqs (4,1) and (4,5) was in principle legitimated by the arbitrariness of v_x in defining $(e^2/\Delta x)/\Delta\varepsilon = \chi_s$ with χ_s expectedly of the order of the unity. Exploit now eq (2,2), for simplicity regarded again at the first order only

$$n_1 = \frac{\delta n}{\delta \log(\Delta\eta')}, \quad \delta \log(\Delta\eta') = \log(\Delta\varepsilon_{n_1+\delta n}) - \log(\Delta\varepsilon_{n_1}),$$

introducing two further energy uncertainty ranges $\Delta\varepsilon_o$ and $\Delta\varepsilon$ whose sizes are by definition intermediate between that of $\Delta\varepsilon_{n_1}$ and that of $\Delta\varepsilon_{n_1+\delta n}$, i.e. $\Delta\varepsilon_{n_1} \le \Delta\varepsilon_o < \Delta\varepsilon \le \Delta\varepsilon_{n_1+\delta n}$. Hence eq (2,2) rewritten as a function of these new ranges takes the form

$$\gamma n_1 = \frac{\zeta\delta n}{\delta \log(\Delta\eta)}, \quad \gamma = \gamma(\Delta\eta), \quad \zeta = \zeta(\Delta\eta), \qquad \delta \log(\Delta\eta) = \log(\Delta\varepsilon) - \log(\Delta\varepsilon_o). \tag{5,5}$$

Now $\Delta\varepsilon_o$ plays the role of fixed reference energy range, likewise as the early $\Delta\varepsilon_{n_1}$ did. The correction coefficients γ and ζ account for the fact that n_1 and $\delta n = n_2 - n_1$ were early defined for $\Delta\varepsilon_o \equiv \Delta\varepsilon_{n_1}$ and $\Delta\varepsilon \equiv \Delta\varepsilon_{n_1+\delta n}$, being therefore $\gamma = 1$ and $\zeta = 1$; having changed the ranges at right hand sides, clearly γ and ζ must be replaced here by γn_1 and $\zeta\delta n$ with $\gamma \neq 1$ and $\zeta \neq 1$, whence their definitions of functions of $\Delta\varepsilon$ once having fixed $\Delta\varepsilon_o$. So the previous eq (2,2) becomes a particular case of the present result (5,5), which reads now

$$\beta(\Delta\eta) = \frac{\delta g}{\delta \log(\Delta\eta)}, \quad \beta(\Delta\eta) = \gamma n_1,$$

$$\delta g = \zeta n_2 - \zeta n_1 = \delta(\zeta n). \tag{5,6}$$

The third equation is interesting as it defines the new range δg. Let the function ζ be someway proportional to $\Delta\eta$, i.e. let ζ decrease with $\Delta\eta$; also, consider the particular case where $\Delta\eta$ is so small that the notation δg can replaced by the familiar differential symbol dg whatever the actual δn might be. Being the range sizes arbitrary, this position about δg is not a hypothesis; it focuses the attention on a particular chance of $\Delta\eta$ that must be taken into account simply because it is allowed and thus to be actually expected. Since a smaller and smaller uncertainty range identifies better and better a local value of the random variable included by its boundaries, $\delta \log(\Delta\eta)$ tends to $d \log(\eta)$; hence the former equation (5,6) tends to the known beta function $\beta(\eta) = dg/d \log(\eta)$ defining the coupling constant g at the energy scale defined by η. This particular limit case helps thus to understand the physical meaning of the ratio in the first eq (5,6), merely written as a function of ranges instead of local values. It is clear the interest to take now $\Delta\eta$ comparable with $\Delta\varepsilon$ of eq (4,1) and (4,5) in order to infer from $\beta(\Delta\eta)$ the function $g(\Delta\eta) \equiv \chi_s$ to be compared with the respective χ_{em} and χ_w. The next task is to calculate the first eq (5,6) in order to confirm that χ_s is of the order of the unity. To this purpose let us expand β in series of powers of δg, i.e. $\beta = \beta_o + \beta_1\delta g + \beta_2\delta g^2 + \cdots$: the coefficient β_o must be equal to zero because of eqs (5,6), whereas $\beta_1 = 0$ as well to fulfill the reasonable condition $\partial\beta/\partial(\delta g) = 0$ of minimum β for $\delta g = 0$. Hence $\beta = \beta_2\delta g^2$, neglecting the higher order terms, requires $\delta g = (\beta_2\delta \log(\Delta\eta))^{-1}$; this appears replacing $1/(\delta \log(\Delta\eta))$ in eq (5,6), which indeed turns into $\beta(\Delta\eta) = \beta_2(\delta g)^2$. According to the fourth eq (5,5), $\delta g = \beta_2/(\log(\Delta\varepsilon/\Delta\varepsilon_o))$ is reducible to the well known form

$$\delta g = \frac{\xi}{\zeta \log(\Delta\eta^2/\Delta\varepsilon_o^2)}, \quad \frac{2\zeta}{\xi} = \beta_2, \quad \Delta\varepsilon_o \approx 0.2 \text{ GeV}. \tag{5,7}$$

The order of magnitude of $\Delta\varepsilon_o$ is easily justified recalling the eq (2,5) of section 2 and the conclusions thereafter inferred: $\Delta\varepsilon_o$ implies that to $\Delta t \approx \hbar/\Delta\varepsilon_o$ corresponds the path $\delta x \approx \hbar c/\Delta\varepsilon_o$ of gluons moving at the light speed to carry the interaction between quarks. The given value of $\Delta\varepsilon_o$ is therefore consistent with the order of magnitude $\delta x \approx 10^{-15}$ m previously quoted for the strong interaction. The result (5,7) and the value of $\Delta\varepsilon_o$ are well known outcomes of quantum chromodynamics; further considerations, in particular about the constants ξ and ζ, are omitted for brevity. This paper aims indeed to show the consistency of the present model based uniquely on eqs (1,1) with the standard features of the strong interactions, not to repeat known concepts.

6 The quark and lepton masses

This section consists of two parts, the first of which concerns the ability of eq (2,4) to describe the ideal masses of isolated quarks. Correlating these masses to the energy ranges $\Lambda_i \equiv \Delta\varepsilon_{n_1+\delta n}$ is in principle sensible first of all regarding the various quarks as a unique class of particles: there would be no reason to expect that different kinds of particles of dissimilar nature are all described by a unique law simply changing a unique distinctive index, here represented by $i \equiv \delta n$. Moreover must hold for the energies of the various quarks a common sort of functional dependence upon δn like that of $\Delta\varepsilon_{n_1+\delta n}$. Eventually, this dependence must still hold even replacing these ranges with the respective average energies $< \varepsilon_{n_1+\delta n} >$ calculated as described in section 4. This last requirement suggests correlating the quark masses with these averages in agreement with the eq (2,4), tanks to the fact that both $< \varepsilon_{n_1+\delta n} >$ and $\Delta\varepsilon_{n_1+\delta n}$ are consistent with their own δn. Indeed an incremental index δn representing the quark energies is defined replacing in eq (2,2) $\log(\Delta\varepsilon_{n_2})$ and $\log(\Delta\varepsilon_{n_1})$ with $\log(< \varepsilon_{n_1} >)$ and $\log(< \varepsilon_{n_2} >)$; a procedure completely analogous yields an equation of the average quantities fully corresponding to eq (2,4). The second point has been explained: the self-interaction of quarks justifies in principle δn

simply admitting that the various quarks are characterized by different self-interaction strengths and thus by distinctively different values of δn. So the critical step is the first one, i.e. whether or not ΔF_x of eq (2,5) really governs the self-interactions of all quarks in order that all of them are related to a unique law (2,4) of δn. This means in practice: (i) regarding one quark delocalized in its own uncertainty range; (ii) thinking that various quarks are characterized by different $\delta(n/m)$ because of their own kind of self-interaction; (iii) assuming that in fact the eq (2,4) accounts for the different numbers of states that characterize uniquely the various quarks. If the functional dependence described by the eq (2,4) is consistent with the three points just mentioned, then Λ_i describes the ideal masses of the quarks as a function of i; also, the point (ii) shows that the energies of this class of particles are really related to their number of allowed states through the self-interaction between quark and gluons.

The estimated masses Q_i of the quarks quoted in literature [14] are reported here:

$$\begin{array}{l} Q_u = 1.7 \leftrightarrow 3.3 \text{ MeV} \\ Q_d = 4.1 \leftrightarrow 5.8 \text{ MeV} \\ Q_s = 80 \leftrightarrow 130 \text{ MeV} \\ Q_c = 1.18 \leftrightarrow 1.34 \text{ GeV} \\ Q_b = 4.13 \leftrightarrow 4.85 \text{ GeV} \\ Q_t = 170.7 \leftrightarrow 173.3 \text{ GeV} \end{array} \qquad (6,1)$$

The mass interval of the "b" quark actually merges two intervals, that reported for the $\overline{MS}$ "mass-independent subtraction scheme" and that of the "1S mass" scheme [14]; the respective mass intervals are $4.19^{+0.18}_{-0.06}$ GeV and $4.67^{+0.18}_{-0.06}$ GeV [15].

It is known that these literature data represent estimates instead of experimental values, as actually isolate quarks do not exist; because of their confinement, the masses are indirectly inferred from scattering experiments. In fact the masses depend on their different combinations in various hadrons and mesons. So the values quoted above must be regarded with carefulness when compared with the results of theoretical calculations. Nevertheless the intervals of values (6,1) do not overlap, which suggests that their order of magnitude is somehow related to and thus at least indicative of the ideal masses of isolated quarks; by consequence it seems also sensible to expect that the sought values of quark masses should fall within these intervals. In lack of further information, therefore, exploit the intervals (6,1) to calculate the average values Q_i:

$$\begin{array}{l} Q_u{}^{(2/3)} = 2.50 \text{ MeV} \\ Q_d{}^{(-1/3)} = 4.95 \text{ MeV} \\ Q_s{}^{(-1/3)} = 105 \text{ MeV} \\ Q_c{}^{(2/3)} = 1.26 \text{ GeV} \\ Q_b{}^{(-1/3)} = 4.49 \text{ GeV} \\ Q_t{}^{(2/3)} = 172 \text{ GeV} \end{array} \qquad (6,2)$$

The superscripts indicate the charges of the respective quarks. These averages have neither specific physical meaning nor come from some particular assumption, they merely represent preliminary starting points defined within realistic intervals; thus their worth is that of reasonable inputs to carry out calculations. The validity of the results inferred in this way relies mostly on their self-consistency; the only initial information is that any sensible output calculated starting from the values (6,2) should expectedly fall within the intervals (6,1). Regard therefore the available data as mere reference values to clarify with the help of eq (2,4) what do Q_i vs i might actually mean in the present context. According to the reasoning carried out in the previous section let us try preliminarily to correlate Q_i with Λ_i putting $\Lambda_i/\Lambda = ((Q_i/U_i)/q)^{1/b}$, where q is a proportionality constant and b a coefficient to be determined by successive calculations; this coefficient fulfills the chance that if $< \Delta\varepsilon_{n_2} > \approx < \Delta\varepsilon_{n_1} >$, i.e. $< \varepsilon_{n_2} > \approx < \varepsilon_{n_1} >$, then the corresponding ratio $(Q_i/U_i q)^{1/b}$ with increasing b anyway matches the limit behavior of Λ_i/Λ whatever q and U_i might be. Initially U_i is justified as mere dimensional factor to be determined; the next results will show that actually it results to be just the potential energy of eq (5,2). Let us sort now the various Q_i by increasing value to check if really the estimated quark masses fulfill the logarithmic dependence of eq (2,4) upon the incremental number of states i, which therefore takes from now on values from 1 to 6. In this way each mass is progressively related to its own increasing i. This expectation is indeed reasonable because $i \equiv \delta n$ defines $\Lambda_i \equiv < \varepsilon_{n_1+\delta n} >$ with respect to a ground reference state number, to which corresponds the reference energy range $\Lambda \equiv < \varepsilon_{n_1} >$. Being by definition $\Lambda_i \equiv \Lambda$ for $\delta n = 0$, one also expects that holds for the eq (2,4) the boundary condition

$$Q_0/U_0 \equiv q \qquad i = 0 \qquad (6,3)$$

whatever b might be; this fact justifies the proposed notation. When handling sets of data, regression calculations are in general needed; the outcomes of these calculations are usually expressed as power series development of an appropriate parameter. Implementing the linear eq (2,4) with the values (6,2) as a function of i, means therefore calculating the best fit coefficients a and b of the form $\log(Q_i/U_i) = a + ib$; clearly n_1 has been included in the regression coefficients. This is easily done regarding Λ_i and Λ of eq (2,4) as follows

$$\log(Q_i/U_i) = a + bi, \qquad a = \log(q), \qquad 1 \le i \le 6. \qquad (6,4)$$

The factor q linking U_i to the reference energy Λ is determined by the boundary condition (6,3); this holds of course even in the presence of higher order terms. The plain first order approximation decided for i agrees with the intent of the present paper: to describe the quarks through an approach as simple as possible and compatible with the minimum amount of input data needed for an unambiguous assessment of re-

sults. So, owing to eqs (5,3) and (5,4), one expects

$$a + bi = \begin{cases} \log(Q_i/U_0) \\ \log\left(Q_i/U_0'(c_i/e)^2\right) \end{cases} \qquad U_0' = \frac{\alpha_s \varepsilon_b e^2 C_j^2}{l_b U_0} \tag{6,5}$$

Now the Δx-dependent behavior of U_i can be checked: if these equations of U_i and the position $\Lambda_i/\Lambda \propto (Q_i/U_i)^{1/b}$ are correct, then both chances (5,3) should somehow appear when exploiting the logarithmic law. A series of plots shows this point step by step starting from the raw data (6,2).

The various Q_i are preliminarily plotted vs i taking all U_i equal to a constant; this first result is reported in fig 1. The boxes represent the input data, the letters between {} identify the quarks, the dot lines describe tentatively their possible connection; the best fit dashed line has a mere indicative meaning of preliminary reference trend. The various points are not completely random, rather they roughly follow an identifiable increase with i. It appears that couples of the various Q_i lie along three lines reasonably parallel each other; so, according to eq (6,4), these lines should be characterized by a unique best fit coefficient b and differ by the coefficient a only. Yet, since each line must be handled in order to fulfill the condition (6,3), the different a are irrelevant: indeed the three regression lines $\log(Q_i) = a_k + bi$, with $k = 1..3$, must be actually plotted as $\log(Q_i/q_k) = bi$ putting $a_k = \log(q_k)$. In effect the fig 2 shows that once having forced the three dotted connections to cross the origin, all quark masses are perfectly aligned along a unique best fit line, whose regression coefficients are: $a_k = 4.7,\ 5.1,\ 5.4$; the respective values of b range between 0.967 and 0.985, i.e. it is reasonably unchanged. Clearly are here concerned the masses of isolated quarks, since the raw data (6,2) have been plotted one by one independently each other. The relevant conclusion is that of having confirmed the validity of eq (2,4) and (2,1): Δx has physical meaning of delocalization range of a unique quark. Considering that the masses spread over 5 orders of magnitude, the result is certainly interesting. If one would calculate the masses of quarks through this plot, however, four constants must be known: three a_k and b: too many, to consider physically meaningful this way of exploiting eq (2,4). The worth of fig 2 is merely heuristic. It must be noted, however, that significant information about b can be obtained through very simple considerations. In the linear regression (6,4), the best fit coefficient b weights the increase of $\log(Q_i)$ as a function of the incremental number of states i. Consider in particular the highest mass Q_6 of the top quark, corresponding to $i = 6$: the greater b, the greater the calculated value of Q_6. So b is expected to be proportional to Q_6. Moreover for the same reason b controls also the masses of lighter quarks for $i < 6$; the link of Q_6 with the masses of all quarks, inherent the plot of fig 2, suggests that the proportionality constant should reasonably have form and physical dimensions somehow related to all quark masses. Put therefore $b = (\sum_{i=1}^{6} Q_i)^{-1} Q_6$,

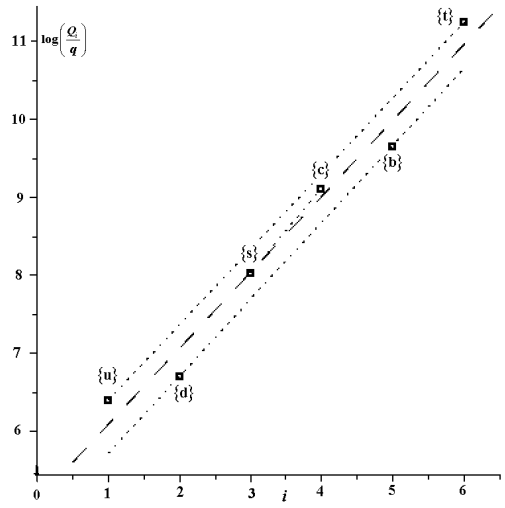

Fig. 2: Plot of $\log(Q_i/q_k)$ vs i; three values of q_k calculated via the boundary condition (6,3) enable a unique trend line of the quark masses with a unique constant U_0.

in which case Q_6 is normalized with respect to the total energy of all possible states allowed between $\Delta\varepsilon_{n_1}$ and $\Delta\varepsilon_{n_1+\delta n}$. Hence the estimates (6,2) yield

$$\frac{Q_6}{\sum_{i=1}^{6} Q_i} = 0.967.$$

In effect, the value of b calculated in this way is very close to that determined in (6,6) via best fit regression.

Yet even three input data to calculate the quarks masses

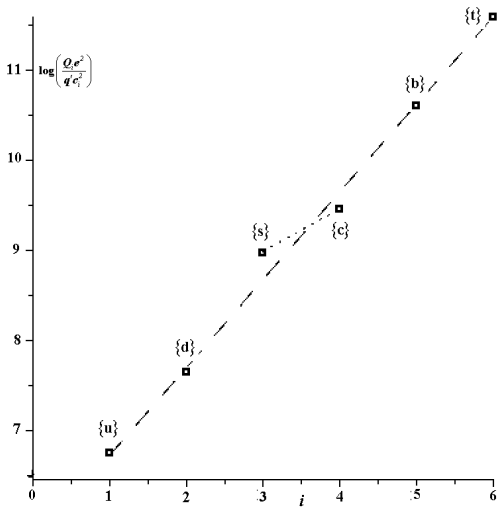

Fig. 4: Plot of $\log(Q_i^*/q')$ vs i with $Q_i^* = Q_i/x_i^2$: here $x_i = const$ for the quarks {c} and {s} and $x_i = c_i/e$ for the other quarks.

Fig. 5: Plot of $\log(Q_i^*/q_k')$ vs i; x_i are defined in fig 4, q_k', with $k = 1, 2$, are calculated in order to fulfil the condition (6,3).

are still too many; certainly there is something else not yet evidenced by the plot of fig 2. Moreover this result, while showing that the idea of concerning the masses of isolated quarks is basically correct, does not highlight anything about the potential energies U_i of eqs (6,5), at the most it could account for U_0 only. Since the idea of considering Q_i/q_k is theoretically too naive, let us regard the various Q_i all together. If so however, despite the previous warnings, the plot of fig 1 is unsatisfactory; owing to the logarithmic ordinate scale, the deviations of the various Q_i from the best fit line are markedly large. Seems however decipherable an unambiguous configuration of these points; this plot prospects the chance of better results. An improved connection between quark masses and i must have exclusively physical valence: here the problem does not concern a random dispersion of experimental measurement errors, but the relationship between masses of isolated quarks and bound quarks on the basis of data extrapolated from the experience; the challenge is to extract the former from the latter trusting to their initial order of magnitude only. The fig 3 reports a new plot where the ratios $(Q_i/U_0)/q$ are replaced by the respective $Q_i e^2/q'c_i^2$, being c_i the electric charges of the various quarks; e is clearly introduced for dimensional reasons. The chance $Q_i e/qc_i$ is not mentioned because found of scarce interest after preliminary checks. From a numerical point of view, therefore, the plain Q_i are now corrected by fractional charge factors $(-1/3)^2$ and $(2/3)^2$. In this way the logarithmic terms are handled exactly as before, which allows the comparison with the former plot: the figure 3 reports again a new best fit line. Now the linear trend of $\log(Q_i e^2/q'c_i^2)$ as a function of i is significantly better than that of fig 1; the {s} and {c} quarks only, both second generation quarks, deviate appreciably from the best fit line; their calculated values consistent with the linear best fit trend are respectively 51 MeV and 1.9 GeV, well outside the literature intervals (6,1). Considering that the orders of magnitude calculated are however globally correct, two chances are in principle admissible: either the literature estimates of the masses of these quarks must be replaced by the values calculated here or some further physical reason, not yet taken into account, enables to modify just these values and align them with the others. The former option is in principle acceptable according to the previous warnings on the literature quark masses, but would conflict with the plot of fig 1: both masses of these quarks were correctly aligned on a similar best fit line before introducing the correction due to their electric charges. So the

latter option seems more stimulating.

Replace therefore $Q_i e^2/q' c_i^2$ of the quarks {c} and {s} only with $Q_i/const$. This idea works well defining *const* appropriately, i.e. in order to fit Q_i of these two quarks to the main best fit line of the other quarks. The fig 4 reports the same data of fig 3, yet replacing e^2/c_i^2 of the quarks {s} and {c} only with a unique value not dependent on c_i; now Q_i^*/q' with $Q_i^* = Q_i/x_i^2$ includes both chances through x_i. The ideal line joining these quark masses is reasonably parallel to the four quark best fit line, i.e. the plot of these two quarks differs trivially from that of the other quarks by the value of the constant a only. As before, in fact this means admitting two values of a: one for the main best fit line, another one for the second generation quark best fit line; of course both values must make the best fit lines compliant with the condition (6,3) via a unique b. The result is shown in fig 5: despite replacing c_i/e with a unique constant is certainly an approximation, nevertheless all quark masses are reasonably represented by a unique eq (2,4). In conclusion, the path from fig 1 to fig 5 was aimed to verify that effectively the logarithmic law (2,4) is expressed via the ratio Q_i/U_i vs the incremental number i of states. The plot of the quark masses Q_i is described by the following equation

$$\begin{array}{lll} \log\left(\dfrac{Q_i^*}{Q_0}\right) = bi & Q_i^* = \dfrac{Q_i}{x_i^2} & b = 0.9723 \\ x_i = c_i/e & Q_0 = 0.556 \text{ MeV} & 1^{st}, 3^{rd} \text{ gen.} \\ x_i = 0.3644 & Q_0 = 1.118 \text{ MeV} & 2^{nd} \text{ generation} \end{array} \qquad (6,6)$$

So b is very similar to that of fig 2. The double value of U_0 corresponds to the two regression constants a allowing to merge the best fit lines of fig 4 according to the condition (6,3); Q_i^* plays the role of an "effective mass" of quarks. The reliability of the results inferred from the plots is assessed recalculating via eqs (6,6) the quark masses and comparing them to the starting values (6,2); one finds $2.32,\ 5.44,\ 1.22 \times 10^2,\ 1.14 \times 10^3,\ 4.50 \times 10^3,\ 1.69 \times 10^5$ MeV that agree reasonably with the literature intervals (6,1). As mentioned at the beginning of this section, this is the basic requirement to be fulfilled. To assess this result also note however that the values (6,2) do not have the rank of experimental data, to be necessarily matched as exactly as possible; as stated before, they have a mere indicative meaning of reference values. Hence the conclusion is that the eqs (6,6) yield a sensible result, while having also the merit of verifying the positions (6,5) strictly related to eqs (5,3). But the most interesting remark concerns U_i, which depends explicitly on the charges c_i in the first and third generation of quarks only; in the second generation it does not, which brings to mind the respective limit cases introduced in eqs (5,3) and further emphasized in eqs (6,5). The generations of quarks are indeed described by $\log(Q_i/U_i) = bi$ with U_i defined by the following equations

$$\begin{array}{ll} U_i^{(\mathrm{ii})} = 0.556(c_i/e)^2 \text{ MeV} & 1^{st}, 3^{rd} \text{ generation,} \\ U_i^{(\mathrm{i})} = 0.148 \text{ MeV} & 2^{nd} \text{ generation.} \end{array}$$

The superscripts are assigned to the generations of quarks by comparison with eqs (5,4) and (6,5); so $U_0 = 0.148$ MeV and $U_0' = 0.556$ MeV.

Some further remarks on this result are also useful. The first concerns the plots of figs 2 and 5: despite the former has been obtained from $\log(Q_i/q_k)$ and the latter from $\log(Q_i/U_i)$ that involves the potential energy, both plots look like and fit surprisingly well the logarithmic law (2,4) despite the quark masses spread over 5 orders of magnitude. These plots are not trivial duplicates: it is interesting the fact that Q_i/U_i takes both forms $Q_i(e/c_i)^2$ and $Q_i/const$, while are determined U_0 and U_0'. On the one hand is remarkable the fact of having identified the mass range as the reason that discriminates the chances (i) and (ii) of eqs (6,5): indeed the mass range of the second generation of quarks is well defined with respect to that of the first and third generations. On the other hand, the fact that both chances are merged in the same plot is itself a further fingerprint of the quantum uncertainty, early introduced because of the mere arbitrariness of Δx. The third remark confirms the fact that Δx is not necessarily the distance between two quarks, it can also be the delocalization range of one quark only; the fact that the plot of fig 5 overlaps very well that of fig 2 shows that even isolated quarks must be regarded as self-interacting and that the interaction potential energy between quarks, the well known eq (5,2) is a replica of the self-interaction potential energy. This conclusion, also supported by the fact that the plot of fig 5 is better than that of fig 1 by introducing Q_i/c_i^2 and not Q_i/c_i, explains why eq (5,2) describing the interaction between different quarks holds also for isolated quarks. The fourth remark concerns the values of the constants U_0 and U_0' reported in eqs (6,6), which describe the asymptotic freedom introduced in the previous section.

Note eventually that the considerations hitherto carried out have assumed already known the quark masses; also, in eqs (6,6) appear several constants to be known "a priori" to carry out the calculations. Moreover, the literature estimates (6,1) appear now as values well configured in the frame of eq (2,4) but not directly supported by experimental measurements. In this respect, a sound proof of their meaning would be to calculate them contextually to other well known and well determined particle mass. The merit of this first part of the section is to have checked the eqs (5,2) and (5,3) via the logarithmic law of eq (2,4). Yet it is also possible to extend further this idea considering together both lepton and quark masses. Indeed a simple question arises at this point: does the eq (6,4) hold also for the leptons? The fact that quarks and leptons are both fundamental bricks of matter suggests the idea that the eq (6,4) could hold for both classes of particles. Moreover note an interesting coincidence: the number of leptons is 6, like that of the quarks. Is this a mere accident or is there some correlation between each quark and each lepton? The next part of the section will show that considering together both kinds of particles allows obtaining all of their

masses as a consequence of a unique principle.

The literature data on the masses L_i of the 6 leptons are summarized here:

$$\begin{array}{ll} e \to 0.51 \text{ MeV}, & \mu \to 105.66 \text{ MeV}, \\ \tau \to 1776.84 \text{ MeV}, & \nu_e \to \ < 2.2 \text{ eV}, \\ \nu_\mu \to < 170 \text{ KeV}, & \nu_\tau \to \ < 15.5 \text{ MeV}. \end{array} \qquad (6,7)$$

The difficulty of comparing calculated and experimental masses concerns now the neutrinos, because of their very scarce interaction with matter and because the neutrino flavor eigenstates are not the same as the mass eigenstates due to the neutrino oscillations [17]. However, being the masses of electron, muon and tau well known, the strategy to carry out the next calculations is: (i) to assume preliminarily the eq (6,4) for the masses of the leptons; (ii) to fit the masses of the neutrinos to the profile required by the logarithmic law via an appropriate correction factor downscaling their upper limit values (6,7); (iii) to look for a unique best fit calculation including both leptons and quarks; (iv) to infer some conclusion about the physical meaning of such a result.

Since the most important task of this section is to find a correlation between the lepton and quark masses previously determined and to confirm the validity of the previous results, the approach proposed here does not concern directly eq (2,4) rewritten in the form (6,4) $\log(L_i) = a' + b'i$ involving the lepton masses only; rather we start looking since the beginning for a connection between L_i and Q_i. Let us show first of all that such a link actually exists, i.e. that are physically sensible logarithmic laws having the forms $\log\left(Q_i^*\right) \pm \log(L_i)$ with Q_i^* defined in eqs (6,6). From $\log(Q_i^*) = a_Q + bi + ci^2 + \cdot\cdot$ and $\log(L_i) = a_L + b'i + c'i^2 + \cdot\cdot$, with $a_Q = \log(Q_0)$ and $a_L = \log(L_0)$ regression constants, one finds first $\log(Q_i^*) \pm \log(L_i) = a_Q \pm a_L + (b \pm b')i + (c \pm c')i^2 + \cdot\cdot$; the higher powers of i have been skipped for brevity, whereas the dimensional factors Q_0 and L_0 are included in the constants a_Q and a_L as in eq (6,4). The fig 6 evidences that the idea of plotting $\log(Q_i^*) + \log(L_i)$ and $\log(Q_i^*) - \log(L_i)$ vs i is sensible: in fact both curves are reasonably definable through appropriate best fit coefficients. To obtain these plots, the neutrino masses, quoted in literature through the respective upper limits only, have been downscaled to the following values

$$\nu_e = 1.802 \text{ eV}, \nu_\mu = 3481.6 \text{ eV}, \nu_\tau = 1.549 \times 10^7 \text{ eV}. \quad (6,8)$$

Moreover the various L_i have been sorted by increasing mass like the respective Q_i^*. This sorting criterion establishes a one-to one correspondence between leptons and quarks that reads

$$\begin{array}{lcccccc} \textit{leptons} & \nu_e & \nu_\mu & e & \nu_\tau & \mu & \tau \\ & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow \\ \textit{quarks} & u & d & s & c & b & t \end{array} \qquad (6,9)$$

Before commenting this correspondence and confirming the validity of eq (2,4) also for the leptons, let us repeat here preliminarily the reasoning previously carried out for the quarks. Calculate $(\sum_{i=1}^6 L_i)^{-1} L_6$ exploiting the values (6,7) and (6,8); one finds

$$\frac{L_6}{\sum_{i=1}^6 L_i} = 0.935 \qquad \left(\frac{Q_6}{\sum_{i=1}^6 Q_i}\right)^2 = 0.936$$

which shows that the lepton equation is related to that of the quarks. To explain this result assume that the normalized values of L_6 and Q_6 are correlated, i.e. $L_6/\sum_j L_j = b' Q_6/\sum_j Q_j$, being b' a constant; imposing then $b' = b$, in order that also $L_6/\sum_j L_j$ be proportional to b of eq (6,5) for the same aforesaid reasons, one finds the given result. These considerations put a constrain on the best fit coefficients of Q_i and L_i vs i. The fig 6 suggests the reasonable chance of introducing a further arbitrary constant b_o that defines the more general linear combinations $\log\left(Q_i^*\right) \pm b_o \log(L_i) = a''_Q \pm b_o a''_L + b''_\pm i + \cdot\cdot$. Hence, multiplying side by side these equations and collecting the constants at right hand side, it must be also true that

$$(\log(Q_i^*))^2 - b_o^2 (\log(L_i))^2 = a''^2_Q - b_o^2 a''^2_L + \cdot\cdot$$

skipping even the first power of i. In effect the advantage of having introduced the arbitrary coefficient b_o is that it can be defined in order to make even the first order term negligible with respect to the constant term, whence the notation reported here; so, neglecting all powers of i, the right hand side reduces to a constant. The last equation reads thus

$$(\log(Q_i^*))^2 = a + (\log(L_i))^2 b, \quad a = a''^2_Q - b_o^2 a''^2_L, \quad b = b_o^2.$$

Now implement again the input data listed in (6,7), (6,8) and (6,2) to check if this last equation correlates sensibly the sets of leptons and quark masses via two constants a and b

Fig. 6: Plot of $\log(Q_i^*/Q_0) \pm \log(L_i/L_0)$ vs i; Q_0 and L_0 are dimensional best fit constants. Q_0 is defined in eqs (6,6).

only; these constants are clearly best fit coefficients that describe the correspondence (6,9). If the zero order approximation just introduced is correct, then trivial regression calculations should yield a sensible statistical correlation of all masses. The best fit coefficients consistent with the zero order approximation of the last equation are

$$a = 45.49178521, \qquad b = 1.039628847. \tag{6,10}$$

So the best fit equation is

$$\log\left(Q_i^{bf}/x_i^2\right) = \pm\sqrt{a + b(\log(L_i))^2};$$

the notation stresses that Q_i^* of eqs (6,6) are replaced by values Q_i^{bf} of Q_i determined by the regression, while the various x_i are of course still that defined in eq (6,6). This result is readily checked calculating

$$Q_i^{bf} = x_i^2 \times 10^{\pm\sqrt{a+b\left(\log(L_i)\right)^2}} \tag{6,11}$$

via the respective lepton masses L_i listed in (6,7), (6,8) and comparing with Q_i reported in (6,1). Note that, because of the exponentials, the decimal places of the regression coefficients are important to reproduce the results of the following calculations. All of the values calculated with the positive sign in eq (6,11)

$$\begin{array}{ll} Q_u^{bf} = 2.50\times10^6 \text{ eV} & Q_d^{bf} = 4.97\times10^6 \text{ eV} \\ Q_s^{bf} = 1.08\times10^8 \text{ eV} & Q_c^{bf} = 1.22\times10^8 \text{ eV} \\ Q_b^{bf} = 4.45\times10^9 \text{ eV} & Q_t^{bf} = 1.75\times10^{11} \text{ eV} \end{array} \tag{6,12}$$

fit surprisingly well the values (6,2) and, mostly important, fall within the estimated intervals (6,1); it is worth noticing that the agreement is much better than that obtained through eqs (6,6). A further remark in this respect is the following. When carrying out the regression calculations with random input data, have been traced the percent deviations of the resulting values of quark and lepton masses with respect to the respective input values; the best self-consistency was found with the true data; the conclusion is that the regression is not mere calculation procedure, but rather a real physical representation of the masses. This also supports the idea that the average values (6,2) of the estimated intervals (6,1) could have an actual physical meaning. Yet are also allowed the following results calculated with the minus sign

$$\begin{array}{ll} q_u^{bf} = 7.91\times10^{-8} \text{ eV} & q_d^{bf} = 2.48\times10^{-9} \text{ eV} \\ q_s^{bf} = 1.64\times10^{-10} \text{ eV} & q_c^{bf} = 1.45\times10^{-11} \text{ eV} \\ q_b^{bf} = 2.77\times10^{-12} \text{ eV} & q_t^{bf} = 1.13\times10^{-12} \text{ eV} \end{array} \tag{6,13}$$

The former set of energies has a literature check through the estimates (6,1), the latter set does not; yet there is no reason to exclude the values (6,13), whose physical meaning will appear shortly. In the latter case the subscripts have a formal physical meaning only, merely reminiscent of the respective quark masses (6,12); nevertheless, it is possible to show the key role of these further energies for the physics of quarks and leptons.

Any statistical regression concerns by definition whole sets of values; here eq (6,11) correlates all masses of leptons and that of all quarks reported in (6,2) and (6,7), (6,8) according to their representation (6,9). The best fit coefficients (6,10) are therefore the fingerprint of *all* masses. Various simulations have been indeed carried out (i) altering deliberately some selected input values of either set of masses, (ii) altering either whole set of masses and (iii) altering both whole sets of masses by means of arbitrary multiplicative factors to find out how the corresponding results are affected; the results, compared with that of eq (6,11) obtained from true values, confirm of course that anyway the new regression coefficients differ from (6,10). The obvious conclusion is that, for some specific reason, just the quoted coefficients (6,10) identify uniquely the fundamental masses of our universe: a is related to their measure units, as previously explained, b controls instead the link between quarks and lepton masses at increasing values of i. Actually one coefficient only is enough to identify all masses; the other is merely associated to it, being concurrently calculated. Otherwise stated, one could assume as a fundamental assumption one of these coefficients only, the other one results consequently determined by the unique set of quark and lepton masses consistent with the former one. Is clear the importance of understanding the specific physical meaning of the particular couple of coefficients (6,10) able to account for the fundamental masses of our universe as a function of *one* predetermined input. Besides the numerical calculation of these masses, however, it seems reasonable to expect that some physical idea is still hidden in eq (6,11).

To investigate this point consider the following equation

$$q_i^o = x_i^2 \times 10^{\pm\sqrt{a+\left(\log(L_i)\right)^2}} \tag{6,14}$$

inferred from (6,11) leaving unchanged a while replacing instead b with the unity. This equation results formally from $\left(\log(q_i^o)\right)^2 = (\log(L_i))^2 + a$, which is interesting because q_i^o and L_i can be interchanged simply changing the sign of a but not its absolute value. Of course the various q_i^o so defined are no longer quark masses; being still related to the respective true lepton masses L_i, however, also q_i^o are somehow related to Q_i.

It is very significant to regard eqs (6,14) thinking Q_i correlated to L_i, which in turn are correlated to q_i^o via one additive constant a only.

So far the experimental masses of quarks and leptons have been introduced as a matter of fact, thus finding that a unique equation, (6,11), accounts for all of them simply postulating a well defined and unique couple of regression constants. Eq (6,14) adds to this standpoint a new perspective: the existence

of a field whose quanta are related to the q_i^o, as a function of which are first calculated L_i via eq (6,14) and then Q_i via eq (6,11). The number of input data confirms that highlighted before, i.e. the quoted value of a only; the masses of both quarks and leptons appear then as consequences of a unique kind of particles, just the q_i^o, since the only possible regression of L_i with Q_i consistent with the given a is that with the concurrent value b. This explains why q_i^o have been defined keeping a and changing b only; even without appearing explicitly appearing in eq (6,14), we know that the latter is required to be just that consistent with the former.

Note now that also eq (6,14) allows two sets of of values, q_i^{o+} and q_i^{o-}, defined by either possible sign of the exponential; it is easy to realize that, likewise as the values (6,12) and (6,13), also now from a numerical point of view $q_i^{o+} >> q_i^{o-}$. This appears regarding all q_i^o together: the resulting total energies corresponding to the positive and negative signs are $\sum_{i=1}^{6} q_i^{o+} = 1.29 \times 10^{11}$ eV and $\sum_{i=1}^{6} q_i^{o-} = 8.189 \times 10^{-8}$ eV. Define therefore the linear combination $q_i^{o+} - q_i^{o-}$ and sum together all i-th terms; one obtains again a total energy

$$\varepsilon_H = 129 \text{ GeV}.$$

Regardless of the numerical values, however, the physical meaning of each term $q_i^{o+} - q_i^{o-}$ is profoundly different from that of the terms q_i^{o+} and q_i^{o-} regarded separately: the masses m_i, charges c_i, spins s_i, colors C_i and so on of these virtual particles, expectedly the same for q_i^{o+} and q_i^{o-} whatever they might be as a consequence of eq (6,14), subtract each other and thus do no longer appear in $q_i^{o+} - q_i^{o-}$. This point is easily highlighted and explained. Actually the eq (6,14) establishes the numerical values of the new energies q_i^{o+} and q_i^{o-}, not their specific forms about which nothing has been hypothesized or is known. The most natural way to regard these quantities, in full line with the basic ideas of the present model, is to relate the various q_i^o to appropriate energy uncertainty ranges as done in eq (2,4); this means assuming for instance

$$q_i^{o+} = \varepsilon_i^+(m_i, c_i, s_i, C_i, ..) - \varepsilon_i^+(0, 0, 0, 0, ..)$$
$$q_i^{o-} = \varepsilon_i^-(m_i, c_i, s_i, C_i, ..) - \varepsilon_i^-(0, 0, 0, 0, ..)$$

with

$$\varepsilon_i^-(m_i, c_i, s_i, C_i, ..) \approx \varepsilon_i^-(0, 0, 0, 0, ..)$$

as well. As repeatedly stressed, both boundaries of any uncertainty ranges are arbitrary. Here we are interested to consider in particular ranges fulfilling the following condition about the upper boundaries:

$$\varepsilon_i^+(m_i, c_i, s_i, C_i, ..) = \varepsilon_i^-(m_i, c_i, s_i, C_i, ..).$$

These positions agree with $q_i^{o+} >> q_i^{o-}$ and also yield

$$q_i^{o+} - q_i^{o-} = \varepsilon_i^-(0, 0, 0, 0, ..) - \varepsilon_i^+(0, 0, 0, 0, ..)$$

that defines $q_i^{o+} - q_i^{o-}$ as the energy uncertainty range of a massless, spinless, chargeless, colorless,.. virtual particle, having in particular boson character. So, when summing up all these terms one finds a total boson energy having the value just quoted. This peculiar energy that accounts for the lepton and quark masses corresponds to a *composite* particle consisting of the sum of 6 terms $q_i^{o+} - q_i^{o-}$ rather than to a truly elementary particle. This conclusion is supported by the fact that the lifetime Δt_H of such a particle should reasonably result from that of its longest life constituent term with $i = 1$, i.e. $\Delta t_H = \hbar/(q_1^{o+} - q_1^{o-})$; one calculates in this way via eq (6,14)

$$q_1^{o+} - q_1^{o-} = 2.50 \text{ MeV}, \qquad \Delta t_H = 2.63 \times 10^{-22} \text{ s}.$$

These last results are reasonable and fully agree with the outcomes of recent experimental measurements.

7 The quantum statistical distributions

This section investigates further consequences of eq (2,2). This part of the paper is thus significant because just this equation leads to eq (2,4), which has been heavily involved to infer the asymptotic freedom equation (5,7) of quarks and the masses of quarks and leptons; confirming once more eq (2,2) means therefore to correlate these results to another fundamental topic of quantum physics concerned in the present section, i.e. the statistical distributions of quantum particles. Eqs (1,1) link the energy range $\Delta\varepsilon$ including the possible energies of a quantum system to its number n of allowed states: the change of energy range size $\delta\Delta\varepsilon = (\hbar/\Delta t)\delta n$ during a given time rage Δt has been concerned in section 2 to calculate the related change $\delta n = n_2 - n_1$ of n, thus obtaining eq (2,4). In that case n_1 was regarded as a fixed quantity, i.e. as a reference number of states as a function of which to define δn. Now we generalize these ideas: both n_1 and n_2 are allowed to change in a quantum system characterized by an initial number of states n_o. If so $\hbar/\Delta t$ can be identically rewritten as $\hbar/\Delta t = \Delta\varepsilon_{n_1}/n_1$ or $\hbar/\Delta t = \Delta\varepsilon_{n_2}/n_2$, because both right hand sides are equivalent reference states in defining δn. So, being both chances alike as well, it is reasonable to expect that $\hbar/\Delta t \propto K_t/(n_1 n_2)$ with $K_t = K_t(\Delta t)$ proportionality factor having physical dimensions of an energy. This position is possible in principle because Δt is arbitrary; so, whatever n_1 and n_2 might be, certainly exists a time length $\Delta t = \Delta t(n_1, n_2)$ that fulfills the proposed correlation. From a formal point of view, assume that $\Delta\varepsilon/n$ of the system is described during Δt by the linear combination $a_1\Delta\varepsilon_{n_1}/n_1 + a_2\Delta\varepsilon_{n_2}/n_2$, being a_1 and a_2 appropriate time dependent coefficients; if so, then $K_t = a_1 n_2 \Delta\varepsilon_{n_1} + a_2 n_1 \Delta\varepsilon_{n_2}$ is defined just by the equation $\delta\Delta\varepsilon/\delta n = \hbar/\Delta t = K_t/(n_1 n_2)$. Since all quantities at right hand side are arbitrary, for simplicity let us approach the problem in the particular case where K_t is regarded as a constant in the following. This chance is obviously also obtainable defining appropriately a_1 or a_2 or both during Δt. The following discussion will show that even this particular case is far reaching and deserves attention.

Write $n_2 = n_o \pm j$ and $n_1 = \pm j$, being n_o a reference fixed number of states and j a variable integer accounting for the change of n_1 and n_2; of course both n_o and j are arbitrary and independent each other, which yields indeed $n_2 - n_1 = n_o$ or $n_2 - n_1 = n_o \pm 2\delta j$ depending on the signs of j. In this way it is possible to describe a steady system with its n_o initial states or an evolving system where is allowed a new number $n' \neq n_o$ of states; since now both n_1 and n_2 are allowed to change, $\delta n = \pm 2\delta j$. Simplifying the notations, the equation inferred from $\delta\Delta\varepsilon/\delta n = K_t/(n_1 n_2)$ of interest for the following discussion reads

$$\frac{\delta\Delta\varepsilon_j}{\delta j} = \frac{2K}{j(n_o \pm j)}, \qquad \delta j = 1,\ 2,.. \tag{7,1}$$

where K must be intended as the constant replacing K_t previously introduced; it is allowed to take both signs, which is avoids writing explicitly $\pm\delta j$. The notation $\Delta\varepsilon_j$ emphasizes the variable number of states appearing at right hand side. To proceed on, consider the case where both j and n_o are large enough to regard approximately the former as a continuous variable, so that $\delta j << j$; so the left hand side can be handled, for mere computational purposes only, as $d\Delta\varepsilon_j/dj$; hence $\Delta\varepsilon_j$ calculated solving the differential equation, results to be

$$\Delta\varepsilon_j = (K'\varepsilon_o/n_o)\log(n_o/j \pm 1) + const, \quad 2K = -K'\varepsilon_o, \tag{7,2}$$

being *const* the integration constant; K' is an arbitrary dimensionless constant and ε_o an arbitrary constant energy. Consider now two boundary conditions of eq (7,2) concerning the respective limit cases (i) $n_o << j$ and (ii) $n_o >> j$. From a mathematical point of view, note that eq (7,2) is obtained by integration of eq (7,1) with respect to j regardless of n_o; hence one could think the cases (i) and (ii) as due to fixed integration limits on dj for two different values of n_o consistent with either inequality, of course without modifying the result of the integration and the subsequent considerations.

In the case (i) holds $n_o/j + 1$ only; putting $const = 0$ and expanding in series the logarithmic term, the right hand side of eq (7,2) reads

$$\begin{gathered} \Delta\varepsilon_j = \frac{w_j K'\varepsilon_o}{j}, \\ w_j = 1 - \frac{n_o}{2j} + \frac{n_o^2}{3j^2} - \ldots, \quad 0 < w_j < 1. \end{gathered} \tag{7,3}$$

Let j be defined between two arbitrary numbers of states j_1 and $j_2 > j_1$; moreover define now K' in order that the sum of all terms $K'w_j$ introduced in the last equation over all values of j fulfills the following condition

$$j_1 \leq j \leq j_2, \qquad \pi_j = K'w_j, \qquad K'\sum_{j_1}^{j_2} w_j = \sum_{j_1}^{j_2} \pi_j = 1;$$

then the result is

$$\pi_j = \frac{j\Delta\varepsilon_j}{\sum_{j=j_1}^{j_2} j\Delta\varepsilon_j}, \quad \varepsilon_o = \sum_{j=j_1}^{j_2} j\Delta\varepsilon_j, \quad \frac{n_o}{j_1} << 1. \tag{7,4}$$

The inequality ensures that is fulfilled the initial condition of the case (i) concerned here, whereas the first eq (7,4) shows the probabilistic character of π_j resulting from the previous positions.

Consider now the limit case (ii). Despite the second eq (7,3) requires in principle a very large number of series terms to express $n_o/j >> 1$, even tending to infinity, there is no reason to exclude that the second equation (7,4) defining $j\Delta\varepsilon_j$ still holds: being K' arbitrary, it can be still defined in order to fulfill the inequality $K'\Sigma_j(1 - n_o/2j + n_o^2/3j^2 + ..) < 1$ whatever the ratio n_o/j might be. On the one hand this inequality can be accepted in principle even though the series consists of an infinite number of terms; in fact the series does not need to be explicitly computed, which makes plausible also the position $\pi_j = K'w_j$. On the other hand, however, in this way the result $j\Delta\varepsilon_j = K'\varepsilon_o w_j$ is not explicitly inferred: the left hand side of the last inequality is indeed undefined. Otherwise stated, without the straightforward hint coming from the case (i) the eqs (7,4) could have been hypothesized only and then still introduced in the case (ii) as plausible inputs but without explanation. Actually, the assessment of the limit case (i) and the subsequent considerations on $w_j K'$ are the points really significant of the present reasoning: while extending the physical meaning of π_j and $j\Delta\varepsilon_j$ also to the case (ii), they ensure the compatibility of the limit cases (i) and (ii). Once again, the arbitrariness of the numbers of states plays a key role to carry out the reasoning.

Looking back to eq (7,2) and multiplying by j both sides, let us write

$$j\Delta\varepsilon_j = K'\varepsilon_o(j/n_o)\log(n_o/j \pm 1) + const\, j. \tag{7,5}$$

According to eqs (7,4) $j\Delta\varepsilon_j/K'\varepsilon_o = w_j$; so, neglecting 1 with respect to n_o/j in agreement with the present limit case (ii) and summing all terms w_j, eq (7,5) yields

$$W = -\sum_{j=j_1}^{j_2}\left(\frac{j}{n_o}\right)\log\left(\frac{j}{n_o}\right) - \sigma\frac{const}{K'\varepsilon_o}, \qquad \sigma = \sum_{j=j_1}^{j_2} j. \tag{7,6}$$

It is useful now to rewrite eq (7,6) as a function of a new variable ξ_j

$$W = -q\sum_{j=j_1}^{j_2}\xi_j\log(\xi_j), \quad const = -\frac{K'\varepsilon_o}{n_o}\log(q), \quad \xi_j = \frac{j}{n_o q},$$

where q is a proportionality factor not dependent on j; it has been defined according to the second equation to eliminate the second constant addend of eq (7,6). The next step is to define j, so far simply introduced as an arbitrary integer without any hypothesis on its actual values, in order that W has specific physical meaning with reference to a thermodynamic system characterized by a number s of freedom degrees. To this purpose assume that j can take selected values n^s only, with n arbitrary integer. This is certainly possible:

nothing hinders calculating the eq (7,2) as a function of n_o/n^s instead of any j progressively increasing; in this way also the eq (7,6) accordingly calculated takes a specific physical meaning consistent with that of the ratios n^s/n_o. Clearly this does not mean trivially renaming j: now n^s reads $\Delta x\Delta p/\hbar^s$, where $\Delta x = \Delta x_1 \cdot \cdot \Delta x_s$ and $\Delta p = \Delta p_1 \cdot \cdot \Delta p_s$. Since therefore $\Delta x\Delta p$ symbolizes a volume in a s-dimensional phase space, $\Delta x\Delta p/\hbar^s$ represents the number of states allowed in this volume. It is known that this ratio introduces the statistical formulation of the entropy [16]; so putting $const/K'\varepsilon_o$ proportional to a new quantity S_0, one finds

$$S = -q\sum_{n=n_1}^{n_2}\xi_n\log(\xi_n), \qquad S_0 = -q\log(\Omega),$$
$$\frac{const}{K'\varepsilon_o} = \frac{1}{\zeta}\frac{S_0}{q}, \qquad \Omega = q^{\zeta/n_o}. \tag{7,7}$$

The notation of the first sum emphasizes that now j takes values corresponding to the possible n^s. The constant of eq (7,6) has been therefore related in the last equation to S_0. The second equation can be regarded as a particular case of the former when the thermodynamic probabilities ξ_j are all equal; while in eq (7,2) j was an arbitrary number progressively increasing fro j_1 to j_2, in eq (7,7) its relationship to n^s does not exclude the chance of coincident values for equal volumes of phase space. It is well known that the results so far exposed introduce the statistical definition of entropy a trivial proportionality factor apart. Note that this result has been obtained in a very different context [12], i.e. to show the quantum character of the Fick diffusion laws as a consequence of eqs (1,1) only; despite the different topic, the theoretical frame is however exactly the same as that hitherto concerned.

Let us return now to the early eq (7,2). Define as usual the energy range as $\Delta\varepsilon_j = \varepsilon'' - \varepsilon'$, so that the eq (7,2) reads $n_o(const + \varepsilon' - \varepsilon'')/K = \log(n_o/j \pm 1)$. Exploit once again the fact that in general the boundary values of the uncertainty ranges are arbitrary; hence, whatever the sign and values of K and $const$ might be, the left hand side can be rewritten as $(\varepsilon_j - \varepsilon_o)/K$, being of course both ε_j and ε_o still arbitrary. So the number of states j of the eq (7,2) reads

$$j = \frac{n_o}{\exp((\varepsilon_j - \varepsilon_o)/K) \mp 1}, \quad \Delta\varepsilon = \varepsilon_j - \varepsilon_o = n_o(const + \varepsilon' - \varepsilon'').$$

The second equation reports again the starting point from which is inferred the former equation to emphasize that, despite the arbitrariness of the boundary values that define the size of the energy uncertainty range, the specific problem determines the values of physical interest. For instance in eq (2,6) has been inferred the Planck law identifying $\Delta\varepsilon_j$ with $h\Delta\nu_j$; clearly the number of states therein appearing is to be identified here with j, whereas n_o can be taken equal to 1 because the photons are bosons. Here the upper sign requires signs of K and $\varepsilon_j - \varepsilon_o$ such that $(\varepsilon_j - \varepsilon_o)/K > 0$ because the number of states j must be obviously positive; instead the lower sign allows in principle both $\varepsilon_o < \varepsilon_j$ and $\varepsilon_o > \varepsilon_j$, as in effect it is well known. To understand these conclusions, let us exploit the reasonable idea that the number j of states allowed for a quantum system is related to the number N of particles of the system. Recall another result previously obtained exploiting eqs (1,1) [7]: half-integer spin particles can occupy one quantum state only, whereas one quantum state can be occupied by an arbitrary number of integer spin particles. In the former case therefore j is directly related to N, i.e. $j = N$ and $n_o = 1$, in the latter case instead in general $N >> j$ without a specific link between j and N. Yet the arbitrariness of n_o makes j suitable to represent any N also in this case as $N = \sum j = n_o \sum(\exp(\Delta\varepsilon_j/K) - 1)^{-1}$. In the classical case where $\Delta\varepsilon_j >> K$, this equation is the well known partition function.

8 Discussion

After the early papers concerning non-relativistic quantum physics [5,6], the perspective of the eqs (1,1) was extended to the special and general relativity; the gravitational interaction was indeed inferred as a corollary just in the present theoretical frame. The problem of examining more in general also other possible forms of quantum interaction appeared next as a natural extension of these results. This paper aimed indeed to infer some basic concepts on the fundamental interactions possible in nature. Even without ambition of completeness and exhaustiveness, the chance of finding some outstanding features unambiguously typical of the electromagnetic, weak and strong interactions has the heuristic value of confirming the fundamental character of eqs (1,1): seems indeed significant that the weird peculiarities of the quantum world are directly related not only to the physical properties of the elementary particles but also to that of their fundamental interactions, which are described in a unique conceptual frame including also the gravity and the Maxwell equations [7]. Now also the gravitational coupling constant, so far not explicitly concerned, is inferred within the proposed conceptual frame. The starting point is again the eq (2,7) rewritten as follows

$$v'_x = -\frac{\Delta F_x \Delta x^2}{n\hbar}, \qquad v'_x = \frac{d\Delta x}{d\Delta t}, \qquad \Delta F_x = F_x - F_{ox}. \tag{8,1}$$

By means of this equation the paper [7] has emphasized the quantum nature of the gravity force, approximately found equal to $\Delta F_x = Gm_a m_b/\Delta x^2$ for two particles of mass m_a and m_b; also, the time dependence of p_x or p_{ox} of $\Delta p_x = p_x - p_{ox}$ was alternatively introduced to infer the equivalence principle of relativity as a corollary. In the present paper, instead, both boundary values of the momentum component range have been concurrently regarded as time dependent to infer the expected potential energy (5,2) of the strong interactions: the reasoning is in principle identical, although merely carried out in a more general way; the form of eq (5,2) comes putting in eq (2,7) both $\dot{p}_x \neq 0$ and $\dot{p}_{ox} \neq 0$, which is the

generalization of the relativistic reasoning carried out in [7]. In fact the eq (2,7), straightforward consequence of eqs (1,1) and thus valid in general, has been reported also in the present paper to better understand these results through its underlying reasoning: what changes is the way it can be exploited to describe specific physical problems, as it has been also emphasized about the physical meaning of v'_x. Now we are interested to implement a particular case of eq (2,7), i.e. the Coulomb law quoted in eq (2,8). The procedure followed below does not need any additional hypothesis with respect to these considerations: it is enough to specify appropriately ΔF_x in eq (8,1).

Consider first the eq (2,8): in the particular case $e' = e$ it yields the Coulomb law $F_x - F_{ox} = \Delta F_x = \pm e^2/\Delta x^2$. Replace this expression into eq (8,1), which reads then

$$v'_x = \pm e^2/n\hbar. \tag{8,2}$$

The ± sign is a trivial feature of the velocity component v'_x along the arbitrary x-axis, it is in fact of scarce interest for the purposes of the present discussion. More interesting is the fact that putting $v'_x = (\alpha/n)c$, as done to infer eq (2,9), one obtains the identity $\alpha/n = e^2/n\hbar c$. This result supports the idea that v'_x/c of eq (8,2) effectively represents a coupling constant: it reads α/n, just the electromagnetic coupling constant found in eq (1,4).

Consider now the gravity force $\Delta F_x = Gm_a m_b/\Delta x^2$ and replace this expression into eq (8,1): so $v'_x = Gm_a m_b/n\hbar$. Comparing this result with the case of the electric force propagating between charged masses, one finds

$$\alpha_G = v'_x/c = Gm_a m_b/n\hbar c. \tag{8,3}$$

Is obvious the reason why the gravitational coupling constant, recognizable at the right hand side, has been formally obtained through elementary considerations identical to that of eq (8,2): the unique eq (8,1) turns into either result simply depending on whether one replaces ΔF_x with $e^2/\Delta x^2$ or $Gm_a m_b/\Delta x^2$. Eqs (8,2) and (8,3) suggest that the gravitational and electromagnetic field propagate at the same rate c: as emphasized when discussing the physical meaning of v_x and v'_x in section 2, the latter is the deformation rate of the space-time range Δx that determines ΔF_x, whereas is instead v_x the real propagation rate of the respective messenger particles in the interaction space-time range Δx; in both cases $\Delta x/\Delta t = c$.

These results are not end points, they have heuristic character. Let us start from eq (8,3) considering for simplicity $m_a = m_b = m$, so that $m = m_P\sqrt{n\alpha_G}$; i.e. any m is proportional to the Planck mass, the proportionality factor being just $\sqrt{n\alpha_G}$. Owing to the small values of α_G, one expects that large values of n are required to fit even small masses. Although α_G depends in general on the specific values of the masses, it is interesting to examine its minimum value corresponding to the particular case where both m_a and m_b represent the lightest elementary particle, the electron neutrino. As concerns the ratio $m_{\nu e}/m_P$ note that $m_{\nu e}$ is a real particle, m_P is a mere definition; so for the former only holds the idea that any particle confined in an arbitrary uncertainty range Δx is characterized in principle by a momentum component gap $\Delta p_x = p_x^{conf} - p_x^{\infty}$ with respect to an ideal unconfined state, see eq (2,1). For the reasoning is irrelevant how an electron neutrino could be confined in practice, because Δx is arbitrary; it could even be the full diameter of the whole universe. It is instead significant in principle that, as already shown in section 4 about the weak interaction boson vectors, it is possible to write for the electron neutrino a delocalization energy $\Delta\varepsilon_{\nu e} = \Delta p_x^2/2m_{\nu e}$ valid for any real object; this reasoning has been in effect exploited in eq (4,13). These considerations aim to conclude that, whatever Δp_x might be, the equation

$$m_{\nu e} = \Delta p_x^2/2\Delta\varepsilon_{\nu e} \qquad \Delta\varepsilon_{\nu e} = m_{\nu e}c^2 \tag{8,4}$$

suggests $m_{\nu e}$ proportional to a reciprocal energy range $\Delta\varepsilon_{\nu e}$ that in turn should be proportional to c^2. If this reasoning is physically sensible, then $m_{\nu e}/m_P \propto c^{-2}$ suggests by consequence $m_{\nu e}/m_P \propto \alpha^2$; since the fine structure constant is proportional itself to c^{-1}, this position simply means including $e^2/\hbar$ into the proportionality constant. Write therefore

$$m_{\nu e}/m_P = \alpha^2/N$$

having called $1/N$ the proportionality constant. The ratio at left hand side is immediately calculated with the help of the first value (6,8), it results equal to 1.5×10^{-28}; the factor $\alpha^2 \approx 5.3\times10^{-5}$ calculates N equal to 3.5×10^{23}, a value surprisingly similar to well known $N = 6.02\times10^{23}$ for the ratio at right hand side. The agreement between these values is really unexpected: while the position $m_{\nu e}/m_P \propto \alpha^2$ could be acceptable at least in principle, is really difficult to understand what the Avogadro number has to do with the present problem. A reasonable idea is to regard α^2/N, perhaps a mere numerical accident, as a whole factor between ordinary mass units and Planck mass units. To support this statement replace in eq (8,4) $\Delta\varepsilon_{\nu e}$ with $m_{\nu e}c^2$, regarded as the average of the boundary values of $\Delta\varepsilon_{\nu e}$; for the following order of magnitude estimate this replacement is acceptable. So, recalling that $\Delta p_x^2 = (n\hbar/\Delta x)^2$ and that actually to calculate $\Delta\varepsilon_{\nu e}$ one should consider $\Delta p_x^2 + \Delta p_y^2 + \Delta p_z^2$, eq (8,4) reads $\Delta x = n\hbar c\sqrt{3/2}/m_{\nu e}c^2$; putting $n = 1$, one finds $\Delta x = 1.3\times10^{-7}$m. Replace now $\Delta\varepsilon_{\nu e}$ with $(N/\alpha^2)\Delta\varepsilon_{\nu e}$: the factor previously found to convert m_{ν_e} into Planck mass units should now convert the energy $\varepsilon_{\nu e}$ from the ordinary units into Plank energy units. Indeed $\Delta x = n\hbar c\sqrt{3/2}\alpha^2/Nm_{\nu e}c^2$ calculated again with $n = 1$ results equal to 1.1×10^{-35}m, which is reasonably comparable with the Planck length $l_P = 1.6\times10^{-35}$m. Actually this result could be expected, because it is based on regarding the energy $\Delta\varepsilon_{\nu e} = \Delta p_x^2/2m_{\nu e}$ as $\Delta\varepsilon_{\nu e} = \Delta p_x^2c^2/2\Delta\varepsilon_{\nu e}$, as already done in section 4; accordingly, this means identifying $\Delta\varepsilon_{\nu e}$ calculated from the confinement uncertainty equation with the mass $m_{\nu e}$

of the particle itself via the factor c^2. This idea was found reasonable to calculate the characteristic length of the weak interaction, eq (4,14), and appears adequate also here because it shows that the conversion factor of m_{ve} into m_P also converts ε_{ve} into E_P.

The main reason for having proposed this result is to stimulate (i) further considerations on the link between α and $\alpha_G^{(ve)}$ and (ii) a greater attention to N when searching fundamental relationships between the constants of nature. Another numerical accident, which is worth noticing here because perhaps of possible interest, concerns the key coefficients (6,10); indeed $\pi a/b = 137.469$, which differs from 137.036 by about 0.3% only. It has been remarked the obvious fact that even small deviations of any lepton or quark mass from the input values (6,2) and (6,7), (6,8) affect the regression coefficients (6,10). So, at least from a numerical point of view, it is sensible to suppose that a very fine-tuning of some among these input values could match exactly the fine structure constant. This optimization is certainly justified: indeed the electron, muon and tau masses only are experimentally known with a degree of accuracy such to exclude any minimum revision; instead, for the reasons previously remarked, there are ample margins of small adjustment for the neutrino and isolate quark masses implemented in the present calculations. On the one hand, such an effort is physically sensible only guessing a good physical reason to expect that the regression coefficients should be actually related to α; on the other hand is evident the interest to provide such an explanation, wholly physical and not merely numerical, of the coefficients that determine the fundamental masses of our universe.

Some further points are still to be better clarified; they pose several questions, some of which are still unanswered. One of them concerns the correspondence (6,9) between leptons and quarks: is it really mere consequence of the increasing order of their masses, thus a mere definition to exploit eq (2,4), or is it actually due to something else still hidden in the correspondence (6,9) and not yet evidenced? But perhaps the most amazing point is that also the leptons fulfill the eq (2,4) just thanks to this correspondence. In the case of quarks, the dependence of their masses on $i \equiv \delta n$ was tentatively explained through the self-interaction of bare quarks with their own clouds of gluons and the self-interaction between these latter: with reference to eq (2,1), a different interaction strength is related both to a dissimilar n/m and to a dissimilar n/V, thus explaining not only the different m of the various quarks but also the equations (5,3) and (6,5). Yet thereafter also the leptons have been handled through the eq (2,4) simply guessing an analogy of behavior for both kinds of fundamental particles of our universe. But, strictly speaking from a physical point of view, why should the lepton masses depend on δn? On the one side the extension of the eq (2,4) certainly works well, because the well known masses of electron, muon and tau particles fit the proposed scheme; the fact of having included these masses among the results calculated through eq (6,10) supports also the values of the masses not experimentally available. On the other side, however, in lack of a self-interaction mechanism characteristic of the quarks only, the question arises: is justified a similar mechanism for the vacuum polarization around the real charges with formation of virtual particle-antiparticle pairs? Does the interaction between these couples of virtual particles/antiparticles surrogate the self-interaction of the quark-gluon plasma? Work is in advanced progress on these points.

Submitted on: December 19, 2012 / Accepted on: December 27, 2012

References

1. Leonhardt U., Paul H. Measuring the quantum state of light. *Progress in Quantum Electronics*, 1995, v. 19, 89–130.
2. Wigner E.P. On the quantum correction for thermodynamic equilibrium. *Physical Review*, 1932, v. 40, 749–759.
3. Allen R.L., Mills D.W. Signal Analysis: Time, Frequency, Scale, and Structure. Wiley-Interscience, NJ, 2004.
4. Tosto S. Spooky action at a distance or action at a spooky distance? *Progress in Physics*, 2012, v. 1, 11–26.
5. Tosto S. An analysis of states in the phase space: the energy levels of quantum systems. *Il Nuovo Cimento B*, 1996, v. 111, 193–215.
6. Tosto S. An analysis of states in the phase space: the diatomic molecules. *Il Nuovo Cimento D*, 1996, v. 18, 1363–1394.
7. Tosto S. Quantum uncertainty and relativity. *Progress in Physics*, 2012, v. 2, 58–81.
8. Glashow S.L. Partial-symmetries of weak interactions. *Nuclear Physics* 1961, v. 22(4), 579–588.
9. Englert F., Brout R. Broken Symmetry and the Mass of Gauge Vector Mesons. *Physical Review Letters*, 1964, v. 13(9), 321–323.
10. Higgs P.W. Broken Symmetries and the Masses of Gauge Bosons. *Physical Review Letters*, 1964, v. 13(16), 508–509.
11. Guralnik G.S., Hagen C.R., Kibble T.W.B. Global Conservation Laws and Massless Particles. *Physical Review Letters*, 1964, v. 13(20), 585–587.
12. Tosto S. Fundamentals of diffusion for optimized applications, Energia, Ambiente, *Innovazione*, 2012, v. 4–5, 94–107.
13. Pohl R. et al. *Nature*, 2010, v. 466, 213–217.
14. Nakamura K. et al., (Particle Data Group), Quarks, 2010, JPG 37, 075021
15. Griffiths D. Introduction to Elementary Particles, Wiley, 1987
16. Landau L., Lifchits E. Physique Statistique (in French), Editions MIR, Moscow, 1967
17. Gonzalez-Garcia M.C., Maltoni M. Phenomenology with massive Neutrinos. *Physics Reports*, 2008, v. 469, 1–129.

Strain Energy Density in the Elastodynamics of the Spacetime Continuum and the Electromagnetic Field

Pierre A. Millette

University of Ottawa (alumnus), Ottawa, Canada. E-mail: PierreAMillette@alumni.uottawa.ca

We investigate the strain energy density of the spacetime continuum in the Elastodynamics of the Spacetime Continuum by applying continuum mechanical results to strained spacetime. The strain energy density is a scalar. We find that it is separated into two terms: the first one expresses the dilatation energy density (the "mass" longitudinal term) while the second one expresses the distortion energy density (the "massless" transverse term). The quadratic structure of the energy relation of Special Relativity is found to be present in the theory. In addition, we find that the kinetic energy pc is carried by the distortion part of the deformation, while the dilatation part carries only the rest-mass energy. The strain energy density of the electromagnetic energy-momentum stress tensor is calculated. The dilatation energy density (the rest-mass energy density of the photon) is found to be 0 as expected. The transverse distortion energy density is found to include a longitudinal electromagnetic energy flux term, from the Poynting vector, that is massless as it is due to distortion, not dilatation, of the spacetime continuum. However, because this energy flux is along the direction of propagation (i.e. longitudinal), it gives rise to the particle aspect of the electromagnetic field, the photon.

1 Introduction

The Elastodynamics of the Spacetime Continuum (*STCED*) is based on the application of a continuum mechanical approach to the analysis of the spacetime continuum [1–3]. The applied stresses from the energy-momentum stress tensor result in strains in, and the deformation of, the spacetime continuum (*STC*). In this paper, we explore the resulting strain energy per unit volume, that is the strain energy density, resulting from the Elastodynamics of the Spacetime Continuum. We then calculate the strain energy density of the electromagnetic field from the electromagnetic energy-momentum stress tensor.

2 Strain energy density of the spacetime continuum

The strain energy density of the spacetime continuum is a scalar given by [4, see p. 51]

$$\mathcal{E} = \frac{1}{2} T^{\alpha\beta} \varepsilon_{\alpha\beta} \tag{1}$$

where $\varepsilon_{\alpha\beta}$ is the strain tensor and $T^{\alpha\beta}$ is the energy-momentum stress tensor. Introducing the strain and stress deviators from (12) and (15) respectively from Millette [2], this equation becomes

$$\mathcal{E} = \frac{1}{2} \left(t^{\alpha\beta} + t g^{\alpha\beta} \right) \left(e_{\alpha\beta} + e g_{\alpha\beta} \right). \tag{2}$$

Multiplying and using relations $e^{\alpha}{}_{\alpha} = 0$ and $t^{\alpha}{}_{\alpha} = 0$ from the definition of the strain and stress deviators, we obtain

$$\mathcal{E} = \frac{1}{2} \left(4te + t^{\alpha\beta} e_{\alpha\beta} \right). \tag{3}$$

Using (11) from [2] to express the stresses in terms of the strains, this expression becomes

$$\mathcal{E} = \frac{1}{2} \kappa_0 \varepsilon^2 + \mu_0 e^{\alpha\beta} e_{\alpha\beta} \tag{4}$$

where the Lamé elastic constant of the spacetime continuum μ_0 is the shear modulus (the resistance of the continuum to *distortions*) and κ_0 is the bulk modulus (the resistance of the continuum to *dilatations*). Alternatively, again using (11) from [2] to express the strains in terms of the stresses, this expression can be written as

$$\mathcal{E} = \frac{1}{2\kappa_0} t^2 + \frac{1}{4\mu_0} t^{\alpha\beta} t_{\alpha\beta}. \tag{5}$$

3 Physical interpretation of the strain energy density

The strain energy density is separated into two terms: the first one expresses the dilatation energy density (the "mass" longitudinal term) while the second one expresses the distortion energy density (the "massless" transverse term):

$$\mathcal{E} = \mathcal{E}_{\parallel} + \mathcal{E}_{\perp} \tag{6}$$

where

$$\mathcal{E}_{\parallel} = \frac{1}{2} \kappa_0 \varepsilon^2 \equiv \frac{1}{2\kappa_0} t^2 \tag{7}$$

and

$$\mathcal{E}_{\perp} = \mu_0 e^{\alpha\beta} e_{\alpha\beta} \equiv \frac{1}{4\mu_0} t^{\alpha\beta} t_{\alpha\beta}. \tag{8}$$

Using (10) from [2] into (7), we obtain

$$\mathcal{E}_{\parallel} = \frac{1}{32\kappa_0} \left[\rho c^2 \right]^2. \tag{9}$$

The rest-mass energy density divided by the bulk modulus κ_0, and the transverse energy density divided by the shear modulus μ_0, have dimensions of energy density as expected.

Multiplying (5) by $32\kappa_0$ and using (9), we obtain

$$32\,\kappa_0\,\mathcal{E} = \rho^2 c^4 + 8\,\frac{\kappa_0}{\mu_0}\,t^{\alpha\beta}t_{\alpha\beta}. \tag{10}$$

Noting that $t^{\alpha\beta}t_{\alpha\beta}$ is quadratic in structure, we see that this equation is similar to the energy relation of Special Relativity [5, see p. 51] for energy density

$$\hat{E}^2 = \rho^2 c^4 + \hat{p}^{\,2} c^2 \tag{11}$$

where $\hat{E}$ is the total energy density and $\hat{p}$ the momentum density.

The quadratic structure of the energy relation of Special Relativity is thus found to be present in the Elastodynamics of the Spacetime Continuum. Equations (10) and (11) also imply that the kinetic energy pc is carried by the distortion part of the deformation, while the dilatation part carries only the rest mass energy.

This observation is in agreement with photons which are massless ($\mathcal{E}_{\|} = 0$), as will be shown in the next section, but still carry kinetic energy in the transverse electromagnetic wave distortions ($\mathcal{E}_{\perp} = t^{\alpha\beta}t_{\alpha\beta}/4\mu_0$).

4 Electromagnetic strain energy density

The strain energy density of the electromagnetic energy-momentum stress tensor is calculated. Note that Rationalized MKSA or SI (Système International) units are used in this paper as noted previously in [3]. In addition, the electromagnetic permittivity of free space ϵ_{em} and the electromagnetic permeability of free space μ_{em} are written with "em" subscripts as the "0" subscripts are used in the spacetime constants. This allows us to differentiate between μ_{em} and μ_0.

Starting from the symmetric electromagnetic stress tensor [6, see pp. 64–66]

$$\Theta^{\mu\nu} = \frac{1}{\mu_{em}}\left(F^{\mu}{}_{\alpha}F^{\alpha\nu} + \frac{1}{4}\,g^{\mu\nu}F^{\alpha\beta}F_{\alpha\beta}\right) \equiv \sigma^{\mu\nu}, \tag{12}$$

with $g^{\mu\nu} = \eta^{\mu\nu}$ of signature $(+---)$, and the field-strength tensor components [6, see p. 43]

$$F^{\mu\nu} = \begin{pmatrix} 0 & -E_x/c & -E_y/c & -E_z/c \\ E_x/c & 0 & B_z & -B_y \\ E_y/c & -B_z & 0 & B_x \\ E_z/c & B_y & -B_x & 0 \end{pmatrix} \tag{13}$$

and

$$F_{\mu\nu} = \begin{pmatrix} 0 & E_x/c & E_y/c & E_z/c \\ -E_x/c & 0 & B_z & -B_y \\ -E_y/c & -B_z & 0 & B_x \\ -E_z/c & B_y & -B_x & 0 \end{pmatrix}, \tag{14}$$

we obtain [6, see p. 66] [7, see p. 141],

$$\begin{aligned} \sigma^{00} &= \tfrac{1}{2}\left(\epsilon_{em}E^2 + \tfrac{1}{\mu_{em}}B^2\right) = \tfrac{1}{2}\epsilon_{em}\left(E^2 + c^2B^2\right) \\ \sigma^{0j} &= \sigma^{j0} = \tfrac{1}{c\mu_{em}}\,(E\times B)^j = \epsilon_{em}c\,(E\times B)^j = \tfrac{1}{c}S^j \\ \sigma^{jk} &= -\left(\epsilon_{em}E^jE^k + \tfrac{1}{\mu_{em}}B^jB^k\right) + \tfrac{1}{2}\delta^{jk}\left(\epsilon_{em}E^2 + \tfrac{1}{\mu_{em}}B^2\right) \\ &= -\epsilon_{em}\left[\left(E^jE^k + c^2B^jB^k\right) - \tfrac{1}{2}\delta^{jk}\left(E^2 + c^2B^2\right)\right] \end{aligned} \tag{15}$$

where S^j is the Poynting vector, and where we use the notation $\sigma^{\mu\nu} \equiv \Theta^{\mu\nu}$ as a generalization of the σ^{ij} Maxwell stress tensor notation. Hence the electromagnetic stress tensor is given by [6, see p. 66]:

$$\sigma^{\mu\nu} = \begin{pmatrix} \frac{1}{2}\epsilon_{em}\left(E^2+c^2B^2\right) & S_x/c & S_y/c & S_z/c \\ S_x/c & -\sigma_{xx} & -\sigma_{xy} & -\sigma_{xz} \\ S_y/c & -\sigma_{yx} & -\sigma_{yy} & -\sigma_{yz} \\ S_z/c & -\sigma_{zx} & -\sigma_{zy} & -\sigma_{zz} \end{pmatrix}, \tag{16}$$

where σ^{ij} is the Maxwell stress tensor. Using the relation $\sigma_{\alpha\beta} = \eta_{\alpha\mu}\eta_{\beta\nu}\sigma^{\mu\nu}$ to lower the indices of $\sigma^{\mu\nu}$, we obtain

$$\sigma_{\mu\nu} = \begin{pmatrix} \frac{1}{2}\epsilon_{em}\left(E^2+c^2B^2\right) & -S_x/c & -S_y/c & -S_z/c \\ -S_x/c & -\sigma_{xx} & -\sigma_{xy} & -\sigma_{xz} \\ -S_y/c & -\sigma_{yx} & -\sigma_{yy} & -\sigma_{yz} \\ -S_z/c & -\sigma_{zx} & -\sigma_{zy} & -\sigma_{zz} \end{pmatrix}. \tag{17}$$

4.1 Calculation of the longitudinal (mass) term

The mass term is calculated from (7) and (17) of [2]:

$$\mathcal{E}_{\|} = \frac{1}{2\kappa_0}\,t^2 = \frac{1}{32\kappa_0}\,(\sigma^{\alpha}{}_{\alpha})^2. \tag{18}$$

The term $\sigma^{\alpha}{}_{\alpha}$ is calculated from:

$$\begin{aligned} \sigma^{\alpha}{}_{\alpha} &= \eta_{\alpha\beta}\sigma^{\alpha\beta} \\ &= \eta_{\alpha0}\sigma^{\alpha0} + \eta_{\alpha1}\sigma^{\alpha1} + \eta_{\alpha2}\sigma^{\alpha2} + \eta_{\alpha3}\sigma^{\alpha3} \\ &= \eta_{00}\sigma^{00} + \eta_{11}\sigma^{11} + \eta_{22}\sigma^{22} + \eta_{33}\sigma^{33}. \end{aligned} \tag{19}$$

Substituting from (16) and the metric $\eta^{\mu\nu}$ of signature $(+---)$, we obtain:

$$\sigma^{\alpha}{}_{\alpha} = \frac{1}{2}\,\epsilon_{em}\left(E^2 + c^2B^2\right) + \sigma_{xx} + \sigma_{yy} + \sigma_{zz}. \tag{20}$$

Substituting from (15), this expands to:

$$\begin{aligned} \sigma^{\alpha}{}_{\alpha} &= \tfrac{1}{2}\,\epsilon_{em}\left(E^2 + c^2B^2\right) + \epsilon_{em}\left(E_x{}^2 + c^2B_x{}^2\right) + \\ &+\epsilon_{em}\left(E_y{}^2 + c^2B_y{}^2\right) + \epsilon_{em}\left(E_z{}^2 + c^2B_z{}^2\right) - \\ &-\tfrac{3}{2}\,\epsilon_{em}\left(E^2 + c^2B^2\right) \end{aligned} \tag{21}$$

and further,

$$\sigma^{\alpha}{}_{\alpha} = \tfrac{1}{2}\,\epsilon_{em}\left(E^2 + c^2B^2\right) + \epsilon_{em}\left(E^2 + c^2B^2\right) - \\ -\tfrac{3}{2}\,\epsilon_{em}\left(E^2 + c^2B^2\right). \tag{22}$$

Hence

$$\sigma^{\alpha}{}_{\alpha} = 0 \tag{23}$$

and, substituting into (18),

$$\mathcal{E}_{\parallel} = 0 \tag{24}$$

as expected [6, see pp. 64–66]. This derivation thus shows that the rest-mass energy density of the photon is 0.

4.2 Calculation of the transverse (massless) term

The transverse term is calculated from (8), viz.

$$\mathcal{E}_{\perp} = \frac{1}{4\mu_0}\, t^{\alpha\beta}t_{\alpha\beta}. \tag{25}$$

Given that $t = \frac{1}{4}\sigma^{\alpha}{}_{\alpha} = 0$, then $t^{\alpha\beta} = \sigma^{\alpha\beta}$ and the terms $\sigma^{\alpha\beta}\sigma_{\alpha\beta}$ are calculated from the components of the electromagnetic stress tensors of (16) and (17). Substituting for the diagonal elements and making use of the symmetry of the Poynting component terms and of the Maxwell stress tensor terms from (16) and (17), this expands to:

$$\sigma^{\alpha\beta}\sigma_{\alpha\beta} = \tfrac{1}{4}\,\epsilon_{em}^2\left(E^2 + c^2B^2\right)^2 + \\ +\epsilon_{em}^2\left[\left(E_xE_x + c^2B_xB_x\right) - \tfrac{1}{2}\left(E^2 + c^2B^2\right)\right]^2 + \\ +\epsilon_{em}^2\left[\left(E_yE_y + c^2B_yB_y\right) - \tfrac{1}{2}\left(E^2 + c^2B^2\right)\right]^2 + \\ +\epsilon_{em}^2\left[\left(E_zE_z + c^2B_zB_z\right) - \tfrac{1}{2}\left(E^2 + c^2B^2\right)\right]^2 - \\ -2\left(S_x/c\right)^2 - 2\left(S_y/c\right)^2 - 2\left(S_z/c\right)^2 + \\ +2\left(\sigma_{xy}\right)^2 + 2\left(\sigma_{yz}\right)^2 + 2\left(\sigma_{zx}\right)^2. \tag{26}$$

The E-B terms expand to:

$$\text{EBterms} = \epsilon_{em}^2\left[\tfrac{1}{4}\left(E^2 + c^2B^2\right)^2 + \\ +\left(E_x{}^2 + c^2B_x{}^2\right)^2 - \left(E_x{}^2 + c^2B_x{}^2\right)\left(E^2 + c^2B^2\right) + \\ +\left(E_y{}^2 + c^2B_y{}^2\right)^2 - \left(E_y{}^2 + c^2B_y{}^2\right)\left(E^2 + c^2B^2\right) + \\ +\left(E_z{}^2 + c^2B_z{}^2\right)^2 - \left(E_z{}^2 + c^2B_z{}^2\right)\left(E^2 + c^2B^2\right) + \\ +\tfrac{3}{4}\left(E^2 + c^2B^2\right)^2\right]. \tag{27}$$

Simplifying,

$$\text{EBterms} = \epsilon_{em}^2\left[\left(E^2 + c^2B^2\right)^2 - \left(E_x{}^2 + c^2B_x{}^2 + \right.\right. \\ \left. +E_y{}^2 + c^2B_y{}^2 + E_z{}^2 + c^2B_z{}^2\right)\left(E^2 + c^2B^2\right) + \\ +\left(E_x{}^2 + c^2B_x{}^2\right)^2 + \left(E_y{}^2 + c^2B_y{}^2\right)^2 + \\ \left. +\left(E_z{}^2 + c^2B_z{}^2\right)^2\right] \tag{28}$$

which gives

$$\text{EBterms} = \epsilon_{em}^2\left[\left(E^2 + c^2B^2\right)^2 - \left(E^2 + c^2B^2\right)^2 + \\ +\left(E_x{}^2 + c^2B_x{}^2\right)^2 + \left(E_y{}^2 + c^2B_y{}^2\right)^2 + \\ \left. +\left(E_z{}^2 + c^2B_z{}^2\right)^2\right] \tag{29}$$

and finally

$$\text{EBterms} = \epsilon_{em}^2\left[\left(E_x{}^4 + E_y{}^4 + E_z{}^4\right) + \\ +c^4\left(B_x{}^4 + B_y{}^4 + B_z{}^4\right) + \\ +2c^2\left(E_x{}^2B_x{}^2 + E_y{}^2B_y{}^2 + E_z{}^2B_z{}^2\right)\right]. \tag{30}$$

Including the E-B terms in (26), substituting from (15), expanding the Poynting vector and rearranging, we obtain

$$\sigma^{\alpha\beta}\sigma_{\alpha\beta} = \epsilon_{em}^2\left[\left(E_x{}^4 + E_y{}^4 + E_z{}^4\right) + c^4\left(B_x{}^4 + B_y{}^4 + \right.\right. \\ \left. +B_z{}^4\right) + 2c^2\left(E_x{}^2B_x{}^2 + E_y{}^2B_y{}^2 + E_z{}^2B_z{}^2\right)\right] - \\ -2\epsilon_{em}^2c^2\left[\left(E_yB_z - E_zB_y\right)^2 + \left(-E_xB_z + E_zB_x\right)^2 + \right. \\ \left. +\left(E_xB_y - E_yB_x\right)^2\right] + 2\epsilon_{em}^2\left[\left(E_xE_y + c^2B_xB_y\right)^2 + \right. \\ \left. +\left(E_yE_z + c^2B_yB_z\right)^2 + \left(E_zE_x + c^2B_zB_x\right)^2\right]. \tag{31}$$

Expanding the quadratic expressions,

$$\sigma^{\alpha\beta}\sigma_{\alpha\beta} = \epsilon_{em}^2\left[\left(E_x{}^4 + E_y{}^4 + E_z{}^4\right) + c^4\left(B_x{}^4 + B_y{}^4 + \right.\right. \\ \left. +B_z{}^4\right) + 2c^2\left(E_x{}^2B_x{}^2 + E_y{}^2B_y{}^2 + E_z{}^2B_z{}^2\right)\right] - \\ -2\epsilon_{em}^2c^2\left[E_x{}^2B_y{}^2 + E_y{}^2B_z{}^2 + E_z{}^2B_x{}^2 + B_x{}^2E_y{}^2 + \right. \\ +B_y{}^2E_z{}^2 + B_z{}^2E_x{}^2 - 2\left(E_xE_yB_xB_y + E_yE_zB_yB_z + \\ \left. +E_zE_xB_zB_x\right)\right] + 2\epsilon_{em}^2\left[\left(E_x{}^2E_y{}^2 + E_y{}^2E_z{}^2 + \right.\right. \tag{32}$$

$$+E_z{}^2E_x{}^2\Big) + 2c^2\left(E_xE_yB_xB_y + E_yE_zB_yB_z + \right.$$
$$\left.+E_zE_xB_zB_x\right) + c^4\left(B_x{}^2B_y{}^2 + B_y{}^2B_z{}^2 + B_z{}^2B_x{}^2\right)\Big]$$

Grouping the terms in powers of c together,

$$\begin{aligned}\frac{1}{\epsilon_{em}^2}\,\sigma^{\alpha\beta}\sigma_{\alpha\beta} = &\Big[\left(E_x{}^4 + E_y{}^4 + E_z{}^4\right) + 2\left(E_x{}^2E_y{}^2 + \right.\\ &\left.+E_y{}^2E_z{}^2 + E_z{}^2E_x{}^2\right)\Big] + 2c^2\Big[\left(E_x{}^2B_x{}^2 + E_y{}^2B_y{}^2 + \right.\\ &\left.+E_z{}^2B_z{}^2\right) - \left(E_x{}^2B_y{}^2 + E_y{}^2B_z{}^2 + E_z{}^2B_x{}^2 + B_x{}^2E_y{}^2 + \right.\\ &\left.+B_y{}^2E_z{}^2 + B_z{}^2E_x{}^2\right) + 4\left(E_xE_yB_xB_y + E_yE_zB_yB_z + \right.\\ &\left.+E_zE_xB_zB_x\right)\Big] + c^4\Big[\left(B_x{}^4 + B_y{}^4 + B_z{}^4\right) + \\ &+2\left(B_x{}^2B_y{}^2 + B_y{}^2B_z{}^2 + B_z{}^2B_x{}^2\right)\Big].\end{aligned} \tag{33}$$

Simplifying,

$$\begin{aligned}\frac{1}{\epsilon_{em}^2}\,\sigma^{\alpha\beta}\sigma_{\alpha\beta} = &\left(E_x{}^2 + E_y{}^2 + E_z{}^2\right)^2 + \\ &+2c^2\left(E_x{}^2 + E_y{}^2 + E_z{}^2\right)\left(B_x{}^2 + B_y{}^2 + B_z{}^2\right) - \\ &-2c^2\Big[2\Big(E_x{}^2B_y{}^2 + E_y{}^2B_z{}^2 + E_z{}^2B_x{}^2 + \\ &+B_x{}^2E_y{}^2 + B_y{}^2E_z{}^2 + B_z{}^2E_x{}^2\Big) - 4\left(E_xE_yB_xB_y + \right.\\ &\left.+E_yE_zB_yB_z + E_zE_xB_zB_x\right)\Big] + c^4\left(B_x{}^2 + B_y{}^2 + B_z{}^2\right)^2\end{aligned} \tag{34}$$

which is further simplified to

$$\begin{aligned}\frac{1}{\epsilon_{em}^2}\,\sigma^{\alpha\beta}\sigma_{\alpha\beta} = &\left(E^4 + 2c^2E^2B^2 + c^4B^4\right) - \\ &-4c^2\Big[\left(E_yB_z - B_yE_z\right)^2 + (E_zB_x - B_zE_x)^2 + \\ &+\left(E_xB_y - B_xE_y\right)^2\Big].\end{aligned} \tag{35}$$

Making use of the definition of the Poynting vector from (15), we obtain

$$\begin{aligned}\sigma^{\alpha\beta}\sigma_{\alpha\beta} = &\,\epsilon_{em}^2\left(E^2 + c^2B^2\right)^2 - \\ &-4\epsilon_{em}^2c^2\left[(E\times B)_x{}^2 + (E\times B)_y{}^2 + (E\times B)_z{}^2\right]\end{aligned} \tag{36}$$

and finally

$$\sigma^{\alpha\beta}\sigma_{\alpha\beta} = \epsilon_{em}^2\left(E^2 + c^2B^2\right)^2 - \frac{4}{c^2}\left(S_x{}^2 + S_y{}^2 + S_z{}^2\right). \tag{37}$$

Substituting in (25), the transverse term becomes

$$\mathcal{E}_\perp = \frac{1}{4\mu_0}\left[\epsilon_{em}^2\left(E^2 + c^2B^2\right)^2 - \frac{4}{c^2}\,S^2\right] \tag{38}$$

or

$$\mathcal{E}_\perp = \frac{1}{\mu_0}\left[U_{em}{}^2 - \frac{1}{c^2}\,S^2\right] \tag{39}$$

where $U_{em} = \frac{1}{2}\,\epsilon_{em}\,(E^2 + c^2B^2)$ is the electromagnetic field energy density.

4.3 Electromagnetic field strain energy density and the photon

$\boldsymbol{S}$ is the electromagnetic energy flux along the direction of propagation [6, see p. 62]. As noted by Feynman [8, see pp. 27-1–2], local conservation of the electromagnetic field energy can be written as

$$-\frac{\partial U_{em}}{\partial t} = \nabla\cdot\boldsymbol{S}, \tag{40}$$

where the term $\boldsymbol{E}\cdot\boldsymbol{j}$ representing the work done on the matter inside the volume is 0 in the absence of charges (due to the absence of mass [3]). By analogy with the current density four-vector $j^\nu = (c\varrho, \boldsymbol{j})$, where ϱ is the charge density, and $\boldsymbol{j}$ is the current density vector, which obeys a similar conservation relation, we define the Poynting four-vector

$$S^\nu = (cU_{em}, \boldsymbol{S}), \tag{41}$$

where U_{em} is the electromagnetic field energy density, and $\boldsymbol{S}$ is the Poynting vector. Furthermore, as per (40), S^ν satisfies

$$\partial_\nu S^\nu = 0. \tag{42}$$

Using definition (41) in (39), that equation becomes

$$\mathcal{E}_\perp = \frac{1}{\mu_0c^2}\,S_\nu S^\nu. \tag{43}$$

The indefiniteness of the location of the field energy referred to by Feynman [8, see p. 27-6] is thus resolved: the electromagnetic field energy resides in the distortions (transverse displacements) of the spacetime continuum.

Hence the invariant electromagnetic strain energy density is given by

$$\mathcal{E} = \frac{1}{\mu_0c^2}\,S_\nu S^\nu \tag{44}$$

where we have used $\rho = 0$ as per (23). This confirms that S^ν as defined in (41) is a four-vector.

It is surprising that a longitudinal energy flow term is part of the transverse strain energy density i.e. S^2/μ_0c^2 in (39). We note that this term arises from the time-space components of (16) and (17) and can be seen to correspond to the transverse displacements along the *time-space* planes which are folded along the direction of propagation in 3-space as the Poynting vector. The electromagnetic field energy density term $U_{em}{}^2/\mu_0$ and the electromagnetic field energy flux term S^2/μ_0c^2 are thus combined into the transverse strain energy density. The negative sign arises from the signature $(+---)$ of the metric tensor $\eta^{\mu\nu}$.

This longitudinal electromagnetic energy flux is massless as it is due to distortion, not dilatation, of the spacetime continuum. However, because this energy flux is along the direction of propagation (i.e. longitudinal), it gives rise to the particle aspect of the electromagnetic field, the photon. As shown in [9, see pp. 174-5] [10, see p. 58], in the quantum theory of electromagnetic radiation, an intensity operator derived from the Poynting vector has, as expectation value, photons in the direction of propagation.

This implies that the $(pc)^2$ term of the energy relation of Special Relativity needs to be separated into transverse and longitudinal massless terms as follows:

$$\hat{E}^2 = \underbrace{\rho^2 c^4}_{\mathcal{E}_\parallel} + \underbrace{\hat{p}_\parallel^2 c^2 + \hat{p}_\perp^2 c^2}_{massless\ \mathcal{E}_\perp} \tag{45}$$

where $\hat{p}_\parallel$ is the massless longitudinal momentum density. Equation (39) shows that the electromagnetic field energy density term ${U_{em}}^2/\mu_0$ is reduced by the electromagnetic field energy flux term $S^2/\mu_0 c^2$ in the transverse strain energy density, due to photons propagating in the longitudinal direction. Thus the kinetic energy is carried by the distortion part of the deformation, while the dilatation part carries only the rest-mass energy, which in this case is 0.

As shown in (9), (10) and (11), the constant of proportionality to transform energy density squared ($\hat{E}^2$) into strain energy density ($\mathcal{E}$) is $1/(32\kappa_0)$:

$$\mathcal{E}_\parallel = \frac{1}{32\kappa_0}\left[\rho c^2\right]^2 \tag{46}$$

$$\mathcal{E} = \frac{1}{32\kappa_0}\hat{E}^2 \tag{47}$$

$$\mathcal{E}_\perp = \frac{1}{32\kappa_0}\left[\hat{p}_\parallel^2 c^2 + \hat{p}_\perp^2 c^2\right] = \frac{1}{4\mu_0}\, t^{\alpha\beta} t_{\alpha\beta}. \tag{48}$$

Substituting (39) into (48), we obtain

$$\mathcal{E}_\perp = \frac{1}{32\kappa_0}\left[\hat{p}_\parallel^2 c^2 + \hat{p}_\perp^2 c^2\right] = \frac{1}{\mu_0}\left[{U_{em}}^2 - \frac{1}{c^2}S^2\right] \tag{49}$$

and

$$\hat{p}_\parallel^2 c^2 + \hat{p}_\perp^2 c^2 = \frac{32\kappa_0}{\mu_0}\left[{U_{em}}^2 - \frac{1}{c^2}S^2\right] \tag{50}$$

This suggests that

$$\mu_0 = 32\kappa_0, \tag{51}$$

to obtain the relation

$$\hat{p}_\parallel^2 c^2 + \hat{p}_\perp^2 c^2 = {U_{em}}^2 - \frac{1}{c^2}S^2. \tag{52}$$

5 Discussion and conclusion

In this paper, we have analyzed the strain energy density of the spacetime continuum in *STCED* and evaluated it for the electromagnetic stress tensor. We have found that the strain energy density is separated into two terms: the first one expresses the dilatation energy density (the "mass" longitudinal term) while the second one expresses the distortion energy density (the "massless" transverse term). We have found that the quadratic structure of the energy relation of Special Relativity is present in the strain energy density of the Elastodynamics of the Spacetime Continuum. We have also found that the kinetic energy pc is carried by the distortion part of the deformation, while the dilatation part carries only the rest mass energy.

We have calculated the strain energy density of the electromagnetic energy-momentum stress tensor. We have found that the dilatation longitudinal (mass) term of the strain energy density and hence the rest-mass energy density of the photon is 0. We have found that the distortion transverse (massless) term of the strain energy density is a combination of the electromagnetic field energy density term ${U_{em}}^2/\mu_0$ and the electromagnetic field energy flux term $S^2/\mu_0 c^2$, calculated from the Poynting vector. This longitudinal electromagnetic energy flux is massless as it is due to distortion, not dilatation, of the spacetime continuum. However, because this energy flux is along the direction of propagation (i.e. longitudinal), it gives rise to the particle aspect of the electromagnetic field, the photon.

Submitted on January 7, 2013 / Accepted on January 11, 2013

References

1. Millette P.A. On the Decomposition of the Spacetime Metric Tensor and of Tensor Fields in Strained Spacetime. *Progress in Physics*, 2012, v. 4, 5–8.
2. Millette P.A. The Elastodynamics of the Spacetime Continuum as a Framework for Strained Spacetime. *Progress in Physics*, 2013, v. 1, 55–59.
3. Millette P.A. Derivation of Electromagnetism from the Elastodynamics of the Spacetime Continuum. *Progress in Physics*, 2013, v. 2, 12–15.
4. Flügge W. Tensor Analysis and Continuum Mechanics. Springer-Verlag, New York, 1972.
5. Lawden D.F. Tensor Calculus and Relativity. Methuen & Co, London, 1971.
6. Charap J.M. Covariant Electrodynamics, A Concise Guide. The John Hopkins University Press, Baltimore, 2011.
7. Misner C.W., Thorne K.S., Wheeler J.A. Gravitation. W.H. Freeman and Company, San Francisco, 1973.
8. Feynman R.P., Leighton R.B., Sands M. Lectures on Physics, Volume II, Mainly Electromagnetism and Matter. Addison-Wesley Publishing Company, Reading, Massachusetts, 1975.
9. Loudon R. The Quantum Theory of Light, Third Edition. Oxford University Press, Oxford, 2000.
10. Heitler W. The Quantum Theory of Radiation, Third Edition. Dover Publications, Inc, New York, 1984.

Liquid Metallic Hydrogen III. Intercalation and Lattice Exclusion Versus Gravitational Settling and Their Consequences Relative to Internal Structure, Surface Activity, and Solar Winds in the Sun

Joseph Christophe Robitaille* and Pierre-Marie Robitaille†

*P.O. Box 21025, Columbus, Ohio, 43221.
†Department of Radiology, The Ohio State University, 395 W. 12th Ave, Columbus, Ohio 43210, USA.
robitaille.1@osu.edu

Invocation of a liquid metallic hydrogen model (Robitaille P.M. Liquid Metallic Hydrogen: A Building Block for the Liquid Sun. *Progr. Phys*., 2011, v. 3, 60–74; Robitaille P.M. Liquid Metallic Hydrogen II: A Critical Assessment of Current and Primordial Helium Levels in Sun. *Progr. Phys*., 2013, v. 2, 35–47) brings with it a set of advantages for understanding solar physics which will always remain unavailable to the gaseous models. Liquids characteristically act as solvents and incorporate solutes within their often fleeting structural matrix. They possess widely varying solubility products and often reject the solute altogether. In that case, the solute becomes immiscible. "Lattice exclusion" can be invoked for atoms which attempt to incorporate themselves into liquid metallic hydrogen. In order to conserve the integrity of its conduction bands, it is anticipated that a graphite-like metallic hydrogen lattice should not permit incorporation of other elements into its in-plane hexagonal hydrogen framework. Based on the physics observed in the intercalation compounds of graphite, non-hydrogen atoms within liquid metallic hydrogen could reside between adjacent hexagonal proton planes. Consequently, the forces associated with solubility products and associated lattice exclusion envisioned in liquid metallic hydrogen for solutes would restrict gravitational settling. The hexagonal metallic hydrogen layered lattice could provide a powerful driving force for excluding heavier elements from the solar body. Herein lies a new exfoliative force to drive both surface activity (flares, coronal mass ejections, prominences) and solar winds with serious consequences relative to the p–p reaction and CNO cycle in the Sun. At the same time, the idea that non-hydrogen atomic nuclei can exist between layers of metallic hydrogen leads to a fascinating array of possibilities with respect to nucleosynthesis. Powerful parallels can be drawn to the intercalation compounds of graphite and their exfoliative forces. In this context, solar winds and activity provide evidence that the lattice of the Sun is not only excluding, but expelling helium and higher elements from the solar body. Finally, exfoliative forces could provide new mechanisms to help understand the creation of planets, satellites, red giants, and even supernova.

Science is a living thing, not a dead dogma. It follows that no idea should be suppressed. That I totally disagree with what you say, but will defend to the death your right to say it, must be our underlying principle. And it applies to ideas that look like nonsense. We must not forget that some of the best ideas seemed like nonsense at first. The truth will prevail in the end. Nonsense will fall of its own weight, by a sort of intellectual law of gravitation. If we bat it about, we shall only keep an error in the air a little longer. And a new truth will go into orbit.

Cecilia Payne-Gaposchkin [1, p. 233]

1 Introduction

As humanity will always be unable to conduct experiments on the stars, insight into stellar physics can only be gained in four steps: 1) the phase of the solar body must be properly ascertained from observational evidence, 2) the substance of which it is comprised must be identified, 3) stellar data must be acquired, and 4) the properties of earthly materials, whose physics might provide at least some level of understanding relative to astrophysical questions, must be taken into account. While such an approach cannot be assured of definitive conclusions, it can nonetheless provide a framework through which the stars can be "understood". Within this context, solar and stellar observations become paramount, as they alone can offer the necessary clues to build realistic models of the stars. Astrophysical data forms the proper foundation for any mathematical treatment. Devoid of observation, theory lacks guidance and leads to stellar models stripped of physical reality.

The postulate that the solar body exists in a liquid state [2, 3] has substantial implications with respect to internal structure and photospheric activity. To understand how the presence of layered graphite-like liquid metallic hydrogen [2, 3]

might alter our insight relative to the Sun, one must turn towards condensed matter physics and the intriguing phenomena associated with both graphite and liquid metallic hydrogen. The consequences are far reaching, touching upon virtually every aspect of astrophysics and provide an elegant setting through which one can begin to understand the most complex observations. Condensed matter offers many advantages not available to gaseous solar models and numerous facts now support a liquid state [4–20].* For instance, evidence suggests that the solar body and the photosphere are behaving as condensed matter [2, 3, 10, 14, 15, 20]. It is not simply that the photosphere gives the appearance of a surface as a result of opacity changes: it is acting as one [14]. The same can be said of every structural element on the Sun, including sunspots, faculae, and granules [15, 20]. The solar body is also behaving as a liquid in sustaining the oscillations which currently occupy helioseismologists. Seismology is a science of the condensed state [10]. Thus, there can be little doubt that the body of the Sun is condensed matter.

Though Gustav Kirchhoff had promoted the idea that the photosphere was liquid, the prevailing models of the period already focused on the gaseous state [21]. By 1865, condensed matter merely floated on the gaseous solar body [21]. Fragmented liquid or solid surfaces continued to survive as a strange addition to gaseous stars [21], but the idea that they were fully liquid never truly materialized in modern astronomy [21]. Finally, liquid stars were definitively abandoned in the days of Sir James Jeans, their last major advocate [22]. Jeans had been unable to identify a proper structural material for his models [22].

Then, in 1935, Wigner and Huntington proposed that pressurized hydrogen could assume a low energy configuration with graphite-like lattice order (see Fig. 1) [23]. In doing so, they unknowingly provided Jeans with a candidate for the solar substance [2, 3], though it is likely that he remained unaware of their solution's value. A layered graphite-like structure was critical to proper solar modeling, as this lattice configuration was closely linked with the study of thermal emission on Earth [24, 25]. Carbon-based materials, such as graphite and soot, are the closest naturally occurring examples of blackbodies [24, 25]. Consequently, they have continued to be vital in the production of such cavities in the laboratory [24, 25]. Thus, a hydrogen based lattice which could adopt a graphite-like structure provides an interesting framework for assembling the Sun. Wigner and Huntington [23] had endowed astrophysics with the perfect candidate for solar material.

In this work, we wish to briefly highlight some of the astrophysical benefits which accompany a liquid metallic hydrogen [23] model of the Sun [2, 3]. Through the liquid model, not only are features on the solar surface given a proper structural foundation, but the entire set of solar observations becomes easily understood [2, 3, 10, 14, 15, 20]. Unlike the gaseous models and their reliance on magnetic fields to explain all aspects of solar activity, the liquid model can secure answers without recourse to such phenomena. Magnetic fields become an effect, not an underlying cause. At the same time, there are ramifications associated with condensed solar matter, especially with respect to gravitational settling, solar activity, and nucleosynthesis. These should be addressed both in the context of existing gaseous models and of the new liquid models of the stars [2, 3].

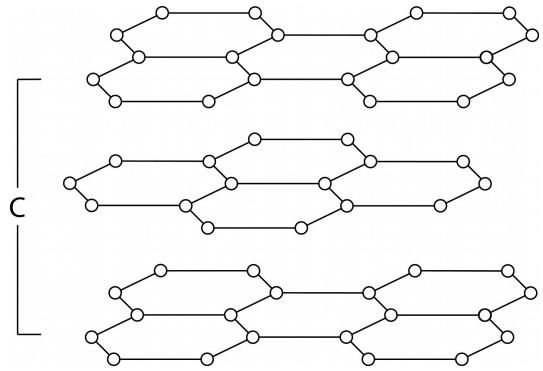

Fig. 1: Schematic representation of the layered hexagonal lattice structure found within graphite and proposed for the liquid metallic hydrogen lattice of the Sun.

2 Solar collapse versus incompressibility

The prevention of solar collapse has always been a central problem with the gaseous models. Theoretical arguments were based on the existence of both gas and radiation pressures in order to balance the masses of the stars against the forces of gravity. In the days of Arthur Stanley Eddington, radiation pressure was believed to play an important role in preventing solar collapse [26]. Over time, this process became generally restricted to supermassive stars [27, p. 180–186]. Solar collapse was prevented by gas pressure [27, p. 132] and radiation thought to contribute only a tiny fraction of the required forces [27, p. 212].

The idea that gas pressure could exist within a star was awkward. On Earth for instance, the atmosphere can be upheld by gas pressure as the planet has a surface through which gas atoms can build positive pressure. Furthermore, the pressure-volume relationship developed using the ideal gas law implied enclosures and rigid surfaces. It was their presence that gave meaning to gas pressure precisely since a rigid compartment defined the volume of interest. But within gaseous stellar models, there are no surfaces. As such, no mechanism exists for speaking of gas pressure.

In his classic text, Donald Clayton would describe the problem as follows: *"The microscopic source of pressure in a perfect gas is particle bombardment.*[1] *The reflection (or*

*The senior author has provided a complete list of his relevant papers to help facilitate the study of this new model.

absorption) of these particles from a real (or imagined) surface in the gas results in a transfer of momentum to that surface. By Newton's second law ($F = dp/dt$)*, that momentum transfer exerts a force on the surface. The average force per unit area is called the pressure. It is the same mechanical quantity appearing in the statement that the quantity of work performed by the infinitesimal expansion of a contained gas is* $dW = PdV$*. In thermal equilibrium in stellar interiors, the angular distribution of particle momenta is isotropic; i.e., particles are moving with equal probabilities in all directions. When reflected from a surface, those moving normal to the surface will transfer larger amounts of momentum than those that glance off at grazing angles*" [28, p. 79]. Clayton's footnote stated: "*In a nonperfect gas strong forces between the particles will represent an additional source or sink of energy for expansions and will therefore contribute to pressure*" [28, p. 79].

There are two problems with Clayton's argument. First, surfaces do not exist within a gaseous Sun. Secondly, by modeling the stars using the ideal gas law, astronomy was requiring elastic collisions between atoms. Yet, if the collisions are elastic, an atom which is moving towards the interior of the Sun could transfer all of its momentum to another atom, without reversing its own direction towards the exterior. In fact, it would simply propel a stationary atom in the interior further inside the Sun. This principle has been well established in the game of billiards. The cue ball can remain completely stationary upon transferring essentially all of its energy to another ball. It is only when a ball hits the banks of the billiard table, or makes use of spin and frictional forces associated with the table surface itself, that it can reverse its momentum. This explains, in the simplest terms, why gas pressure cannot exist within a gaseous Sun devoid of real surfaces and subject to elastic collisions. No net force can be generated with "imaginary surfaces" as the particles have equal probabilities of moving in all directions and transfer their momentum perfectly with no change of direction. A real surface is required to generate a net directional force and such structures cannot exist within a gaseous Sun. Therefore, modern solar models are unable to prevent internal collapse by resorting to gas pressure. In the absence of sufficient radiative forces, gaseous stars collapse.

At the same time, the use of gas models introduced many complications in astronomy. The first was summarized in Eddington's concern regarding internal heating, as stars became increasingly dense: "*I can hardly see how a star which has once got into this compressed condition is ever going to get out of it. . . Imagine a body continually losing heat but with insufficient energy to grow cold*" [29, p. 172]. Ralph H. Fowler would solve Eddington's dilemma. In 1926 [30], he adapted Fermi-Dirac statistics to stellar problems (e.g. [27, p. 118–128]). Stars could now grow cold. Donald Clayton highlighted the salient aspects of Fowler's solution: "*The physical basis for the resolution of this problem is the thermodynamic peculiarity of a degenerate gas: the temperature no longer corresponds to kinetic energy. The electrons in a zero-temperature degenerate gas must still have large kinetic energy if the density is great*" [28, p. 104]. In fact, Fowler's treatment was so theoretically powerful and the arguments so elegant [30], that gaseous stellar models now dominate astronomy. Nonetheless, no mechanism existed for generating gas pressure within Sun-like stars behaving as ideal gases [27, p. 130–132]. Fowler's solution addressed much later stages of stellar evolution [30].

Conversely, liquids are, by their nature, essentially incompressible. Thus, the problem of solar collapse does not occur within the condensed matter context [2, 3], because the layered graphite-like structure of liquid metallic hydrogen (see Fig. 1) would act to uphold the solar mass. Still, it is anticipated that the hexagonal lattice of metallic hydrogen can become slightly compressed with increasing internal solar pressures. The essentially incompressible nature of liquids implies that, while resisting compression, they remain subject to pressure effects to a small extent. Therefore, it is reasonable to anticipate that liquid metallic hydrogen becomes more metallic farther in the solar interior assuming a Type II lattice [2, 3]. The lower pressures of the photosphere would be conducive to supporting a less dense solar lattice (Type-I) with associated decreased metallicity [2, 3]. Conversely, since the Wilson effect [31] implies that sunspots are depressed relative to the photospheric level, it is reasonable to infer the presence of a Type-II lattice with its increased metallicity in these structures [2, 3]. In addition, as facular material is tightly associated with sunspots and may well have been ejected from such regions, it was not unreasonable to extrapolate that their increased metallicity occurs as a result of assuming a Type-II lattice, despite the fact that they appear to float on the photospheric surface [20].

3 Gravitational Settling Versus Restricted Diffusion

Within the context of the gaseous models [32, 33] atoms and ions can diffuse freely within stellar bodies. At the same time, since certain elements are heavier than others, it could be expected that they would slowly move towards the interior of a star through the action of gravitational settling. In fact, such a concept was advanced to explain the lack of helium lines in certain B type stars [34]. Long before, Henry Russell had minimized the idea that heavy elements were gravitationally settling in the Sun: "*It does not appear necessary, therefore, to assume that downward diffusion depletes the sun's atmosphere of the heavier elements, though the possibility of such an influence remains*" [35, p. 59]. Of course, gravitational settling could potentially invalidate all elemental abundances in stellar atmospheres obtained from spectroscopic lines.

Kippenhahn and Weigert discussed both temperature and pressure diffusion (gravitational settling) in their text on "*Stellar Structure and Evolution*" [27, p. 60–61]. They con-

cluded that temperature diffusion was astrophysically irrelevant in the Sun and that diffusion effects were, in general, important only in "special cases" not including the Sun [27, p. 60–61]. Today, the effect of gravitational settling has been included in the calculation of standard solar models [32, 33]. In part, this was because it improved the agreement with the p-mode oscillations from helioseismology: "*One of the principal improvements that has been made in recent years is to include in the calculations the effects of element diffusion. In the absence of an external field, diffusion smooths out variations. However, in the case of the Sun, the stronger pull of gravity on helium and the heavier elements causes them to slowly diffuse downward (towards the solar interior) relative to hydrogen . . . Models that include at least helium diffusion agree with helioseismological determinations of the depth of the convective zone, while neglecting diffusion entirely leads to disagreement with the helioseismological data*" [33]. Gravitational settling was embraced; for gaseous models had no other means of accounting for helioseismological observations.

Within a liquid metallic hydrogen model of the Sun, the free diffusion of the elements becomes highly restricted, as the layered lattice structure of the solar body acts to inhibit the flow of atoms. Rapid diffusion of elements should occur primarily in the layers between the hexagonal liquid metallic hydrogen planes. Such motion may be facilitated by lattice distortions in the hexagonal hydrogen planes in a manner similar to that observed in graphite intercalation compounds.

4 Intercalation and Graphite

Graphite [36–38] can be made to interact with various reagents such that non-carbon atoms occupy lattice points between the hexagonal carbon planes forming intercalation compounds [39–43]. Layered intercalation compounds (see Fig. 2) are created when intraplanar binding forces are much stronger than interplanar forces: "*The most important structural characteristic of graphite intercalation compounds is the occurrence of separate graphite and intercalate layers due to the very strong intraplanar binding and the weak interplanar binding. Thus, the graphite layers retain the basic properties of pristine graphite, and the intercalate layers behave similarly to the parent intercalate material*" [39, p. 36].

In the graphite case, the hexagonal plane excludes non-carbon atoms, the intercalant. In doing so, intercalant atoms can profoundly alter the electrical, thermal, magnetic properties of graphite by acting as electron donors (i.e. Li, K), or acceptors (i.e. $FeCl_3$, HF, BF_3), to the hexagonal plane [39–43]. As a result, graphite intercalation compounds can range from superconductors to insulators [39] with their conductivity often exceeding that of classic metals [43, p. 190]. They consequently occupy an important place in solid state physics. Graphite intercalation compounds can also undergo phase transitions including "*changes in interlayer ordering and changes in intralayer or in-plane ordering, magnetic transitions, and superconductive transition. Structural phase transitions have been induced by variation of the temperature, pressure, and in some cases by variation of the vapour pressure of the intercalant*" [39, p. 55–56]. The presence of intercalated atoms can weaken the interlayer attractive forces within graphite. Since the concentrations of the intercalate can be varied, it is possible to build intercalation compounds wherein many adjacent graphite layers are interrupted by the occasional intercalate layer (see Fig. 3). The stage index, *n*, characterizes the number of graphite layers between intercalation layers (e.g. [39] and [43, p. 88]). In the laboratory, *n* usually ranges from 1 to 10 [39].

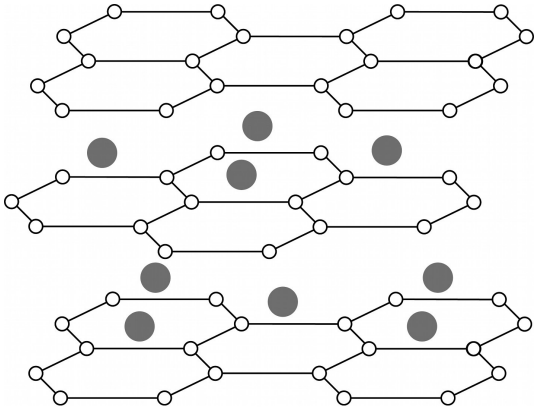

Fig. 2: Schematic representation of an intercalation compound. Non-carbon elements are located between layers of pristine graphite.

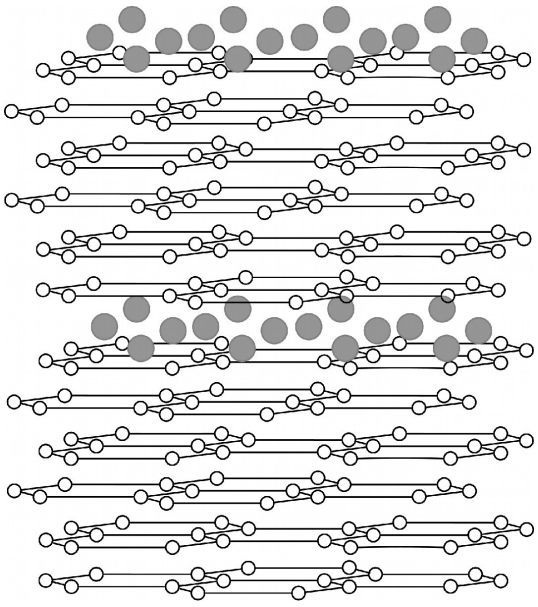

Fig. 3: Schematic representation of the stage index, n, in an intercalate compound, where n = 6.

Graphite intercalation compounds are known to relieve internal strains by undergoing exfoliation [39, p. 9] whereby a great expansion along the c-axis (see Fig. 1) occurs usually due to elevated temperatures [44]. The temperature required for exfoliation is linearly dependent on applied load against the sample [44]. Higher breakaway temperatures, or temperatures of exfoliation, are required under increased pressure. Expansions of the c-axis lattice dimensions of up to a factor of 300 have been reported [44]. These can be violent, even explosive, events wherein layers of material can be torn away from the underlying structure (see e.g. [39, p. 9] and [43, p. 403–413]). They occur as a result of gases being expelled from the graphite intercalated compound. The resultant products are characterized as "*spongy, foamy, low-density, high-surface-area carbon materials*" [43, p. 403].

Martin and Broklehurst [44] performed detailed studies of exfoliation which involved the effect of "*restraining loads on suppressing the onset of exfoliation*" [43, p. 406]. Enoki et al. describe the situation as follows: "*According to [Martin and Broklehurst's] model, the intercalate undergoes a phase change to the vapor phase, forming disk-shaped bubbles of radius* r *and height* I_c *in the interlayer region between graphite planes, with gas pockets accumulating in certain regions where diffusion is facilitated by the presence of defects. Exfoliation then occurs when the gas pressure exceeds the internal stress parallel to the c-axis*" [43, p. 406]. Expressions for the forces involved can be derived, assuming the ideal gas law [44].

Lattice exclusion remains the central lesson of these experiments: the graphite hexagonal planes continue to exclude the intercalate and struggle to remain "*pristine*" even at the cost of exfoliation. Such behavior has strong ramifications when considering the graphite-like liquid metallic hydrogen lattice believed to exist within the Sun [2, 3].

5 Intercalation and Stellar Matter

Graphite's tendency to remain pristine and exclude other elements from its hexagonal plane, even through the process of exfoliation, has important consequences for solar physics. Thermal emission arguments have led Robitaille [2] to postulate that liquid metallic hydrogen in the Sun must adopt a graphite-like layered arrangement. Should this be correct, then liquid metallic hydrogen should be excluding other elements from its hexagonal plane and constantly working to drive them out of the solar body. Such lattice exclusion and the possibility that stars might undergo processes like exfoliation could play a crucial role in at least five separate aspects of solar and stellar dynamics: 1) supplying the driving forces for solar winds, 2) generating the settings for flares, coronal mass ejections, and prominences, 3) accounting for the eleven year solar cycle, 4) providing an alternative explanation for planet and satellite formation, and 5) explaining the existence of red giants and supernovae. Each of these areas could consume many years of study as the liquid metallic hydrogen model of the Sun is adopted. Suffice it, for now, to address these briefly.

5.1 Solar Winds

In modern gaseous models, magnetic fields are thought to be produced by the flow of isolated charged particles within the solar body. In order to prevent collapse, the Sun remains in perfect hydrostatic equilibrium wherein the forces of gravity are balanced by gas and radiation pressure [27, p. 6–7]. However, the preservation of hydrostatic equilibrium severely limits all proposals advanced for the existence of solar winds. An object in equilibrium cannot easily be driving material away from itself.

Conversely, in a condensed model of the Sun, a layered liquid metallic hydrogen lattice exists (see Fig. 1) which is dominated by hexagonal hydrogen planes [2, 3]. Such a lattice restricts the translation of protons within each hexagonal hydrogen layer while permitting electrons to flow in the associated conduction bands [2]. The ability to create conduction bands provides the interatomic binding forces needed to stabilize the hydrogen framework. Proton-proton distances are restricted in order to establish optimal quantum mechanical conditions for these conduction bands. This alone stabilizes the lattice. Since hydrogen atoms possess a single electron and these are restricted to the conduction bands, no conventional bonding can occur. All elements other than hydrogen would be excluded from the hexagonal layer in order to maintain its structural integrity and electronic structure. Protons could be thought of as constantly working to expel elements from the hexagonal planes. This would severely limit the flow of non-hydrogen elements. Each hydrogen layer would act as a barrier to diffusion along the c-axis (see Fig. 1), while providing a channel for rapid elemental diffusion in the region between two hexagonal layers. Herein can be found the driving force for the solar winds and the variable elemental compositions they present due to solar activity [3].

5.2 Flares, Coronal Mass Ejections, and Prominences

In the gaseous models of the Sun, solar flares and coronal mass ejections are considered to be magnetic phenomena [45–48] and are produced by invoking magnetic reconnection [49, 50]. As a gaseous Sun is devoid of a real surface, no other means of generating the required energy is available: "*The magnetic energy stored in the corona is the only plausible source for the energy released during large solar flares. During the last 20 years most theoretical work has concentrated on models which store magnetic energy in the corona in the form of electrical currents, and a major goal of present day research is to understand how these currents are created, and then dissipated during a flare*" [50]. In such a scenario, the corona provides the driving force for expelling atoms from the Sun.

Solar flares are well known to produce helium abundance enhancements (HEA) and have been suggested as the cause of significant ^{3}He HEAs [45]. In an impulsive flare, the ^{3}He/^{4}He ratio can be assumed to approach 1 [51] and thousand-fold enhancements of the ratio have been reported [52] . Solar energetic particle events can result in 100–10,000 fold enhancements of heavy element to oxygen ratios relative to the quiet corona [52]. Solar atmospheric ratios of Mg/O, Si/O, Fe/O and Ne/O can all be substantially elevated with flare activity [51]. In active coronal regions, significant (3–4 fold) elemental enhancements of elements with a first ionization potential (FIP) less than 10 eV can be observed with respect to the quiet photosphere [53, 54]. Within bright active regions, a further twofold elemental enhancement can be detected [55]. The absolute abundance of potassium and calcium are greater in flare plasma than in the photosphere [54].

Magnetic reconnection [49, 50], the physical mechanism invoked to drive solar flares in the gaseous models, cannot easily account for the variable elemental abundances associated with flares and coronal mass ejections [56, 57]. As a parallel, models of quiescent coronal loops result in a 10 fold excess of helium to hydrogen when a 10% helium abundance is assumed for the chromosphere [58]. Such tremendous excesses of helium call for much lower chromospheric helium abundances, but these are incompatible with levels required to account for helium in the solar winds [58]. In addition, in order to explain O and Ne abundances in the fast solar winds, a coronal He abundance of 20–40% is required [59]. The model assumes gravitational settling in the corona [59], which is highly unlikely to take place. As such, the gaseous models are struggling to coherently resolve elemental abundances in the solar winds as a result of the interaction between coronal loops, the chromosphere, and the corona. The situation relative to understanding elemental abundances in flares and coronal mass ejections is equally tenuous.

Long ago, Friedrich Zöllner recognized that solar flares required regions of increased pressure in the solar interior [60]. He placed a liquid layer within his gaseous Sun: "*we must therefore conclude that the layer of division consists of an incandescent liquid*" [60]. The need to generate pressure was justified, but could not easily survive within a fully gaseous solar model.

In the liquid metallic hydrogen model of the Sun, solar flares, coronal mass ejections, and prominences can be explained by the process of intercalation and exfoliation, as described above by Martin and Broklehurst [44]. The pressure anticipated by Zöllner [60] is produced when the intercalate atoms increasingly populate the region between two adjacent hydrogen layers. A rapid increase in temperature in this region, presumably due to localized nuclear reactions (see section 5), generates a gaseous phase whose elevated pressures manifest as solar activity. Therefore, solar flares, coronal mass ejections, and prominences share a common mechanism of formation. Their subtle differences result only from the depth of formation. Magnetic fields are not required to produce these phenomena. They are merely altered by their presence.

5.3 The Eleven Year Solar Cycle

The existence of the eleven year solar cycle remains incompletely understood [61–66]. Nonetheless, increased solar activity is associated with changes in the solar dynamo which characterize the 11 year cycle [61, 64]. Cycle periods as great as 2,400 years have been postulated [66]. Solar inertial motion (SIM), wherein the location of the center of the Sun's mass in the solar system drifts due to interaction with the giant planets [61–66], has been postulated as a possible cause of increased activity. Still, as Cionco and Compgnucci highlight: "*at present there is no clear physical mechanism relating these phenomena*" [64]. How can planetary rotations and the associated *SIM* trigger solar activity? Perhaps the Sun is already predisposed to increased surface turbulence and requires only a simple disturbance to initiate activity. In this regard, insight can be gained from the condensed model of the Sun [2, 3].

In the context of a liquid metallic hydrogen model [2, 3], non-hydrogen elements reside in the layers between hydrogen hexagonal planes forming an intercalate arrangement (see Fig. 2). With solar nuclear activity (see section 5), these interplanar regions become increasingly populated and possible intercalate lattice points occupied. Eventually, localized saturation of a given intercalate layer takes place. The maximal concentration of intercalating atoms has been reached. When this occurs, only slight disturbances, such as found through solar inertial motion, could trigger solar activity and cause the intercalate atoms to be ejected from interior layers. Solar activity then becomes linked to the need to eject saturating levels of non-hydrogen elements from the solar body. As the rate of nuclear activity must remain rather constant over the time frames involved, the Sun is constantly building elements in its interior (see section 5), degassing, and repeating the entire process. The driving force for degassing becomes lattice exclusion, but the trigger to release the instability may, or may not, remain linked to solar inertial motion.

5.4 Planet, Red Giant, and Supernova Formation

The formation of planets around a star presents unique challenges to astronomy. Many ideas have surfaced and are taught in introductory astronomy courses [67, p. 285–290]. With time, Laplace's Nebular Hypothesis [68, 69], initially proposed by Emanuel Swedenborg [70, p. 240–272], evolved into the Solar Nebular Disk Model (*SNDM*) [71]. The latter continues to be the most widely accepted theory for the formation of the solar system [71]. Yet, the problem of planet and satellite formation is far from resolved (e.g. [72–74]). In part, this is because the planets cannot be currently conceived as ejected from a young active gaseous solar mass. The prob-

lem is removed when the Sun becomes condensed matter and exfoliative forces can be harnessed to promote planet formation, especially for the solid planets of the inner solar system. The central requirement appears to be that interlayer elemental abundance must be permitted to increase dramatically in one region of the solar interior, followed by ejection from the hydrogen lattice. Over time, the Sun could thus transfer some of its angular momentum to the planets. A similar approach could be utilized to help explain satellite formation around the giant planets, as they are also rich in hydrogen [75–77].

On a tangential note, exfoliation might well account for the very low density and great dimensions of the red giants, as the experiments of Martin and Broklehurst suggest [44]. A red giant would remain condensed matter in that it was formed through a process of exfoliation from a star which had permitted a nearly uniform stage index to develop in its interior. A trigger finally turned the intercalate rapidly into the gaseous phase resulting in a red giant. In the final expanded star the dimensions would be enormous and the density greatly reduced, despite the preservation of condensed matter for the metallic hydrogen framework. Interlayer gas pressure between the layers of the expanded star would help to maintain its structural integrity. Supernova could be envisioned as produced in a similar manner, but with non-uniform staging in the interior. For instance, a band or core of intercalate material in the precursor star rapidly enters the gas phase and explodes its liquid metallic hydrogen envelope, while compressing its hydrogen core. In the end, the advantages of adopting a liquid metallic hydrogen model for the Sun are numerous and its consequences extend much beyond the solar system.

6 Evolution and Nuclear Reactions in Gaseous Stars

With the publication of the *Origin of Species* [78] Charles Darwin would send shock waves not only throughout the biological sciences, but also in areas seemingly as far removed as astronomy. The great American father of solar astronomy, George Ellery Hale, commented as follows in the first line of his text devoted to stellar evolution and experimental astronomy: "*It is not too much to say that the attitude of scientific investigators towards research has undergone a radical change since the publication of the Origin of Species*" [79]. Hale expanded on this concept throughout his first chapter, as he elegantly intertwined biological evolution and astronomy. Hale also highlighted the conflict which Herbert Spencer [21], the prominent evolutionist, had with the astronomers: "*convinced that the principle of evolution must operate universally, and that the stars must have their origin in the still unformed masses of the nebulae, [Spencer] ventured to question the conclusion that the resolution of nebulae into stars was only a question of resolving power. He had not long to wait . . .* " [79, p. 47].

Given Hale's fame as an observer for first reporting the presence of magnetic fields on the Sun [80], his leadership in constructing four record setting telescopes (at Yerkes (1), Mount Wilson (2), and Palomar (1) [81]), and his role in establishing the *Astrophysical Journal* [82], it is not surprising that *The Study of Stellar Evolution* [79] has profoundly affected the course of modern astrophysics. George Ellery Hale's interest in stellar evolution [28, 83–87] was certain to ascend to a preeminent position in modern astronomy. At the same time, since prolonged biological evolution was also associated with increased functional abilities, astronomers quickly adopted the same concepts relative to the stellar evolution. As stars aged their core temperatures increased and gradually acquired the ability to make heavier elements. Astronomers began to see the stars not only as progressing through a life cycle, but also, as endowed with different synthetic abilities. Older stars possessed hotter cores, and hence, could sustain nuclear processes thought to require higher temperatures – the synthesis of heavier and heavier elements. On the surface at least, the theory was elegant with the exception of one very serious consideration: the gaseous Sun was deprived of the ability to directly synthesize the elements.

Early on, the fathers of stellar nucleosynthesis, such as Gamow [88, 89], Bethe [90–92], von Weizsäcker [93], and Hoyle [94, 95] would advance the idea that helium could be built from hydrogen within the stars. From the onset, nucleosynthesis was linked to stellar evolution [88, 89]. Gamow believed that "*different rates of energy liberation must be due to different physical conditions inside the stars and chiefly to differences in their central temperature*" [83, p. 116]. The p–p reaction [90], which assembled helium directly from proton combinations while relying on positron and neutrino emission, was believed to be active only in low weight main sequence stars [83, p. 118]. However, for stars larger than the Sun much of the synthesis of ^{4}He came from the carbon-nitrogen-oxygen (CNO) cycle which had been independently proposed by Bethe and von Weizsäcker [91–93]. Interestingly, while the cycle required three elements of intermediate weight, Hans Bethe insisted that: "*no element heavier than* ^{4}*He can be built up in ordinary stars*" [92]. He argued, "*The heavier elements found in stars must therefore have existed already when the star was formed*" [92]. With those words, most of the stars were deprived of their ability to make any element beyond helium, despite the fact that mankind would eventually synthesize much heavier elements.

Bethe, of course, based his ideas on the probability of nuclear reactions in the gas phase [92, p. 435]. This was appropriate for gaseous solar models. Reaction energies were derived using accelerators and nucleosynthesis in the stars became strictly dependent on our understanding of reactions in gases. The idea that many particles could be combined simultaneously within a condensed lattice would have greatly lowered the energy required to synthesize the heavier elements. Such a concept was never applied to the Sun. Soon a detailed work by Burbidge et. al [96] organized the entire field into an

elaborate theory of nucleosynthesis which covered all of the elements. This work would continue to influence nucleosynthesis in the stars until the present day [97]. Nonetheless, the Sun itself had been crippled. All of the elements in the solar system, other than helium, had been produced by early generation stars which no longer existed.

7 Nucleosynthesis and Condensed Matter

Perhaps the greatest advantage of the liquid metallic hydrogen model of the Sun rests in the fact that atomic positions become restricted to lattice points and subject to the forces associated both with solar pressures and lattice vibrations. Hydrogen is confined to its hexagonal planes and all other elements to the intercalate positions between the hydrogen planes. The synthesis of helium would be driven by the need to relieve the strains of stellar pressures on the underlying lattice. Two protons combine to form a deuteron, with positron and neutrino emission as in the p–p reaction [98]. Upon formation, the deuteron could immediately combine with another in-plane proton resulting in the formation of ^{3}He, which would be ejected from the lattice plane into the intercalate layer. As p–p reactions continue, the population of ^{3}He would expand, and soon continue to react producing ^{4}He, as expected from branch 1 of the p–p chain [98]. With time, the intercalate region would become the birthplace of all the elements. Pressure and lattice vibrations alone can be viewed as controlling the reactions with protons readily available from the hexagonal plane. All stars gain the ability to synthesize every element [19]. Multiple elements could react simultaneously in the intercalate layer because of lattice vibrations. This greatly lowers the energy requirements on a given species for nuclear reaction. Eventually, as elemental concentrations build, the stresses against the hexagonal hydrogen planes would increase. These could then break and the intercalate region expand beyond the confines of strict lattice points. Intercalation now abandoned in this region, thick layers of non-hydrogen elements could arise. These would continue to act as nuclear furnaces. During periods of increased solar activity, localized changes in temperature could vaporize these areas and release newly synthesized elements to the stellar atmosphere beyond the solar surface. During planet formation, such regions could simply be expelled, with (or perhaps without) vaporization, from the interior of the Sun.

8 Conclusions

Much speculation has been offered in this work and the end result was deliberate. In order to consider the condensed models of the Sun, scientists must ponder upon the ability to explain the highest amount of observable phenomena in a manner consistent with known physics. The great solar physicist John Bahcall once commented: "*Science progresses as a result of the clash between theory and experiment, between speculation and measurement*" [99]. In earlier work, considerable focus was placed on establishing what was known about the Sun and the evidence it displayed with respect to its phase and composition [2–20]. Ample proof supports the idea that the Sun exists in the condensed state and Occam's razor would slice in its favor.

Given the elevated levels of hydrogen in the universe [100], a liquid metallic hydrogen framework appears not only reasonable but, in light of its thermal emission, necessary [2, 3]. The unique link between graphite and the layered form of metallic hydrogen, as first proposed by Wigner and Huntington [23], presents enormous potential to refine our concept of the stars. In this regard, graphite intercalation compounds bring a wealth of behavioral and structural information crucial to understanding the heavens [39–44]. The layered nature of liquid metallic hydrogen [23] would not only support the Sun from collapse, but would also severely limit any gravitational settling. Furthermore, exfoliation in graphite intercalate compounds [44] has profound consequences, regarding stellar structure and behavior. Solar winds and solar activity (flares, coronal mass ejections, prominences) become inherently linked to preserving the hydrogen nature of the Sun [3]. The conversion of intercalated atoms from the liquid to the gas phase, as proposed by Martin and Broklehurst [44], has profound implications towards driving solar activity which will forever remain unavailable to gaseous models. The hypothesis that the solar cycle originates from the degassing of non-hydrogen elements and their expulsion from the interior is unique to the liquid metallic hydrogen model. For the first time, a reasonable thesis is being advanced to explain both solar activity and cycles. A mechanism thereby becomes available to those who believe that solar inertial motion might trigger solar activity [61–66]. In addition, the idea that a layered metallic hydrogen lattice will choose to exclude non-hydrogen elements and sequester them within the Sun could add much needed insight relative to the formation of the planets. Exfoliation of a metallic hydrogen lattice of uniform stage might well account for both the size and density of the red giants. Most importantly, this model enables elemental synthesis in the stars. Hexagonal hydrogen planes harbor the p–p reactions, while the interlayers between proton planes become furnaces of more advanced nuclear synthesis.

There is a great deal to be gained by considering a liquid metallic hydrogen model of the Sun. Yet, in this approach, the solar lattice appears to possess long range order on par with solids, despite its liquid state [18]. Given the dimensions involved on the solar surface, even solids might appear to act as liquids. But nonetheless, the model claims the liquid state as more in keeping with observation. In this respect, the authors emphasize that long range lattice order seems to be preserved in the liquid metallic hydrogen framework of the photosphere and solar body. The Sun is fully behaving as condensed matter. As such, this thesis has been built on observation, in keeping with the philosophy of Cecilia Payne:

"The future of a subject is the product of its past, and the hopes of astrophysics should be implicit in what the science has already achieved. Astrophysics is a young science, however, and is still, to some extent, in a position of choosing its route; it is very much to be desired that present effort should be so directed that the chosen path may lead in a permanently productive direction. The direction in which progress lies will depend on the material available, on the development of theory, and on the trend of thought . . . The future progress of theory is a harder subject for prediction, than the future progress of observation. But one thing is certain: observation must make the way for theory, and only if it does can the science have its greatest productivity . . . There is hope that the high promise of astrophysics may be brought to fruition."

Cecilia Payne-Gaposchkin [1, p. 199–201]

Acknowledgment

Luc Robitaille is acknowledged for the preparation of figures.

Dedication

This work is dedicated to Lindsey Marie Robitaille.

Submitted on: January 6, 2013 / Accepted on: January 10, 2013
First published in online on: February 2, 2013

References

1. Haramundanis K. Cecilia Payne-Gaposchkin: An autobiography and other recollections (2nd Edition), Cambridge University Press, Cambridge, U.K., 1996.
2. Robitaille P.M. Liquid metallic hydrogen: A building block for the liquid Sun. *Progr. Phys.*, 2011, v. 3, 60–74.
3. Robitaille P.M. Liquid Metallic Hydrogen II: A critical assessment of current and primordial helium levels in Sun. *Progr. Phys.*, 2013, v. 2, 35–47.
4. Robitaille P.M. The collapse of the Big Bang and the gaseous Sun. *New York Times*, March 17, 2002, p. A10 (available online: http://thermalphysics.org/pdf/times.pdf).
5. Robitaille P.M. Evidence for a liquid plasma model of the Sun. *Am. Phys. Soc. Meeting — April*, 2004, S280002.
6. Robitaille P.M. The Sun as a hot liquid plasma: additional evidence. *Am. Phys. Soc. Meeting — Ohio Spring*, 2004, S50002.
7. Robitaille P.M. The photosphere as condensed matter. *Am. Phys. Soc. Meeting — Ohio Fall*, 2004, S60005.
8. Robitaille P.M. The Sun as a hot liquid plasma: more evidence. *Am. Phys. Soc. Meeting — NE Fall*, 2004, S10004.
9. Robitaille P.M. The Sun as a high energy/high density liquid metallic hydrogen plasma. *The 33rd IEEE International Conference on Plasma Science*, June 4-8, 2006, Traverse City, Michigan, p. 461, DOI:10.1109/PLASMA.2006.1707334.
10. Robitaille P.M. The solar photosphere: Evidence for condensed matter. *Progr. Phys.*, 2006, v. 2, 17–21 (also found in slightly modified form within *Research Disclosure*, 2006, v. 501, 31–34; title #501019).
11. Robitaille P.M. A high temperature liquid plasma model of the Sun. *Progr. Phys.*, 2007, v. 1, 70–81 (also in arXiv: astro-ph/0410075).
12. Robitaille P.M. A radically different point of view on the CMB. In: *Questions of Modern Cosmology — Galileo's Legacy*, ed. by M. D'Onofrio and C. Burigana, Springer, New York, 2009.
13. Robitaille P.M. Liquid metallic hydrogen: Building block of a liquid Sun. *Am. Phys. Soc. Meeting — Ohio Spring*, 2011, D4.00005.
14. Robitaille P.M. On the Presence of a Distinct Solar Surface: A Reply to Hervé Faye. *Progr. Phys.*, 2011, v. 3, 75–78.
15. Robitaille P.M. On Solar Granulations, Limb Darkening, and Sunspots: Brief Insights in Remembrance of Father Angelo Secchi. *Progr. Phys.*, 2011, v. 3, 79–88.
16. Robitaille P.M. On the Temperature of the Photosphere: Energy Partition in the Sun. *Progr. Phys.*, 2011, v. 3, 89–92.
17. Robitaille P.M. Stellar opacity: The Achilles heel of the gaseous Sun. *Progr. Phys.*, 2011, v. 3, 93–99.
18. Robitaille P.M. Lessons from the Sun. *Progr. Phys.*, 2011, v. 3, 100–102.
19. Robitaille P.M. Nucleosynthesis of the elements and the liquid metallic hydrogen model of the Sun. *Am. Phys. Soc. Meeting — Four Corners Annual*, 2012, D1.00026.
20. Robitaille P.M. Magnetic Fields and Directional Spectral Emissivity in Sunspots and Faculae: Complimentary Evidence of Metallic Behavior on the Surface of the Sun. *Progr. Phys.*, 2013, v. 1, 19–24.
21. Robitaille P.M. A thermodynamic history of the solar constitution — I: The journey to a gaseous Sun. *Progr. Phys.*, 2011, v. 3, 3–25.
22. Robitaille P.M. A thermodynamic history of the solar constitution — II: The theory of a gaseous Sun and Jeans' failed liquid alternative. *Progr. Phys.*, 2011, v. 3, 41–59
23. Wigner E. and Huntington H.B. On the possibility of a metallic modification of hydrogen. *J. Chem. Phys.*, 1935, v. 3, 764–70.
24. Robitaille P.M. Blackbody radiation and the carbon particle. *Progr. Phys.*, 2008, v. 3, 36–55.
25. Robitaille P.M. Kirchhoff's Law of Thermal Emission: 150 years. *Progr. Phys.*, 2009, v. 4, 3–13.
26. Eddington A.S. On the radiative equilibrium of the stars. *Mon. Not. Roy. Astron. Soc.*, 1916, v. 77(1), 16–35 (Also found in Lang K.R. and Gingerich O.: A Source Book in Astronomy and Astrophysics, 1900–1975. Harvard University Press, Cambridge, MA, 1979, p. 225–235).
27. Kippenhahn R. and Weigert A. Stellar structure and evolution. Springer-Verlag, Berlin, 1990.
28. Clayton D.D. Principles of stellar evolution and nucleosynthesis. MgGraw-Hill, New York, 1968.
29. Eddington A.S. The internal constitution of the stars. Cambridge University Press, Cambridge, U.K., 1926.
30. Fowler R.H. On dense matter. *Mon. Not. Roy. Astron. Soc.*, 1926, v. 87, 114–122.
31. Wilson A. Observations on the solar spots. *Phil. Trans. Roy. Soc.*, 1774, v. 64, 1–30.
32. Bahcall J.N. and Pinsonneault M.H. Standard solar models, with and without helium diffusion, and the solar neutrino problem. *Rev. Mod. Phys.*, 1992, v. 64, no.4, 885–926.
33. Bachall J.N., Pinsonneault M.H. and Wasserburg G.J. Solar models with helium and heavy-element diffusion. *Rev. Mod. Phys.*, 1995, v. 67, no. 4, 781–808.
34. Greenstein G.S., Truran J.W. and Cameron A.G.W. Helium deficiency in old halo B type stars. *Nature*, 1967, v. 213, 871–873.
35. Russell H.N. On the composition of the Sun's atmosphere. *Astrophys. J.*, 1929, v. 70, 11–82.
36. Kelly B.T. Physics of graphite. Applied Science Publishers, London, U.K., 1981, p. 34–61.
37. Delhaès P. World of carbon — vol. 1: Graphite and precursors. Gordon and Breach Science Publishers, Amsterdam, The Netherlands, 2001.
38. Pierson H.O. Handbook of carbon, graphite, diamond and fullerenes: Properties, processing and applications. Noyes Publications, Park Ridge, N.J., 1993.

39. Dresselhaus M.S. and Dresselhaus G. Intercalation compounds of graphite. *Adv. Phys.*, 2002, v. 1, no. 1, 1–186 (reprinted from *Adv. Phys.*, 1981, v. 30(2), 139–326).
40. Pietronero L. and Tosatti E. Physics of intercalation compounds. Springer-Verlag, Berlin, 1981.
41. Zabel H. and Solin S.A. Graphite intercalation compounds I: Structure and dynamics. Springer-Verlag, Berlin, 1990.
42. Dresselhaus M.S. and Kalish R. Ion implantation in diamond, graphite and related materials. Springer-Verlag, Berlin, 1992.
43. Enoki T., Suzuki M. and Endo M. Graphite intercalation compounds and applications. Oxford University Press, Oxford, U.K., 2003.
44. Martin W.H. and Brocklehurst J.E. The thermal expansion behavior of pyrolytic graphite-bromine residue compounds. *Carbon*, 1964, v. 1, no. 2, 133–141.
45. Kahler S.W. Solar flares and coronal mass ejections. *Ann. Rev. Astron. Astrophys.*, 1992, v. 30, 113–141.
46. Priest E.R. Solar flare theory and the status of flare understanding. In *High Energy Solar Physics: Anticipating HESSI. ASP Conf. Ser.*, 2000, v. 206, 13–26.
47. Priest E.R. and Forbes T.G. The magnetic nature of solar flares. *Astron. Astrophys. Rev.*, 2002, v. 10, 313–377.
48. Hudson H.S. Global properties of solar flares. *Space Sci. Rev.*, 2011, v. 158, 5–41.
49. Holman G.D. The mysterious origin of solar flares. *Sci. Am.*, 2006, v. 294, no. 4, 38–45.
50. Forbes T.G. Magnetic reconnection in solar flares. *Geophys. Astrophys. Fluid Dynam.*, 1991, v. 62, 15–36.
51. Ramaty R., Mandzhavidze N., Kozlovsky B. and Murphy R.J. Solar atmospheric abundances and energy content in flare-accelerated ions from gamma-ray spectroscopy. *Astrophys. J.*, 1995, v. 455, L193–L196.
52. Reames D.V. and Ng C.K. Heavy-element abundances in solar energetic particle events. *Astrophys. J.*, 2004, v. 610, 510–522.
53. Laming J.M. Non-WKB modles of the first ionization potential effect: Implications for solar coronal heating and the coronal helium and neon abundances. *Astrophys. J.*, 2009, v. 695, 954–969.
54. Doschek G.A., Feldman U. and Seely J.F. Elemental abundances from solar flare spectra. *Mon. Not. Roy. Astron. Soc.*, 1985, v. 217, 317–326.
55. Ciaravella A., Raymond J.C., Li J., Reiser P., Gardner L.D., Ko Y.K. and Fineschi S. Elemental abundances and post-coronal mass ejection current sheet in a very hot active region. *Astrophys. J.*, 2002, v. 575, 1116–1130.
56. Feldman U., Landi E. and Laming J.M. Helium Abundance in High-Temperature Solar Flare Plasmas. *Astrophys. J.*, 2005, v. 619, no.2, 1142–1152.
57. Andretta V., Mauas P.J.D., Falchi A. and Teriaca L. Helium Line Formation and Abundance during a C-Class Flare. *Astrophys. J.*, 2008, v. 681, no. 1, 650–663.
58. Killie M.A., Lie-Svendsen Ø. and Leer E. The helium abundance in quiescent coronal loops. *Astrophys. J. Let.*, v. 632, no. 2, L155–L158.
59. Byhring H.S., Esser R. and Lie-Svendsen Ø. O and Ne in H-He fast solar wind. *Astrophys. J.*, 2011, v. 743, no. 2, 205(11p).
60. Zöllner F. On the temperature and physical constitution of the Sun. *Phil. Mag. 4th Series*, 1870, v. 40, 313–327 (essentially reprinted in: Zöllner F. On the Sun's temperature and physical constitution. *Nature*, 1870, v. 2(52), 522–526).
61. Hathaway D.H. The solar cycle. *Living Rev. Solar Phys.*, 2010, v. 7, 1–66.
62. Schwentek H. and Elling W. A possible relationship between spectral bands in sunspot number and the space-time organization of our planetary system. *Solar Phys.*, 1984, v. 93, no. 2, 403–413.
63. Grandpierre A. On the origin of solar cycle periodicity. *Astrophys. Space Sci.*, 1996, v. 243, no. 2, 393–400.
64. Cioncoa R.G. and Compagnucci R.H. Dynamical characterization of the last prolonged solar minima. *Adv. Space Res.*, 2012, v. 50, no. 10, 1434–1444.
65. Tan B. Multi-timescale solar cycles and the possible implications. *Astrophys. Space Sci.*, 2011, v. 332, no. 1, 65–72.
66. Charvátová I. Can origin of the 2400-year cycle of solar activity be caused by solar inertial motion. *Ann. Geophysicae*, 2000, v. 18, 399–405.
67. Payne-Gaposhchkin C. and Haramundanis K. Introduction to astronomy (2nd Edition), Prentice-Hall Inc., Englewood Cliffs, N.J., 1970.
68. Laplace P.S. Exposition du système du monde. Imprimerie du Cercle-Social, Paris, 1796 (available online: http://dx.doi.org/10.3931/e-rara-497; Also available in English: Pond J. The system of the world, London, 1809).
69. Numbers R.L. Creation by Natural Law: Laplace's Nebular Hypothesis in American Thought. Seattle, 1977, p. 124-132.
70. Swedenborg E. The Principia; or the first principles of natural things, being new attempts towards a philosophical explanation of the elementary world. Translated by: Augustus Clissold, W. Newbery, London, 1846.
71. Woolson M.M. Solar system — Its origin and evolution. *Quarterly J. Roy. Astron. Soc.*, 1993, v. 34, 1–20.
72. Montmerle T., Augerneau J.C., Chaussidon M., Gounell M., Marty B. and Morbidelli A. 3. Solar system formation and early evolution: The first 100 million years. *Earth, Moon, and Planets*, 2006, v. 98, 39–95.
73. Thommes E.W., Duncan M.J. and Levison H.F. The formation of Uranus and Neptune among Jupiter and Saturn. *Astrophys. J.*, 2002, v. 123, 2862–2883.
74. Stevenson D.J. Origin of the moon — The collision hypothesis. *Ann. Rev. Earth Planet Sci.*, 1987, v. , 271–315.
75. Nellis W.J., Ross M. and Holmes N.C. Temperature measurements of shock-compressed hydrogen: Implications for the interior of Jupiter. *Science*, 1995, v. 269, no. 5228, 1249–1252.
76. Nellis W.J., Weir S.T. and Mitchell A.C. Metallization and electrical conductivity of hydrogen in Jupiter. *Science*, 1996, v. 73, no. 5277, 936–938.
77. Vorberger J., Tamblyn I., Militzer B. and Bonev S.A. Hydrogen helium mixtures in the interior of giant planets. *Phys. Rev. B*, 2007, v. 75, 024206(1–11).
78. Darwin C. On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life. John Murray, London, U.K. 1859.
79. Hale G.E. The study of stellar evolution: An account of some recent methods of astrophysical research, The decennial publications of the University of Chicago — Second Series, Vol. X. University of Chicago Press, Chicago, I.L., 1908.
80. Hale G. E. On the probable existence of a magnetic field in sun-spots. *Astrophys. J.*, 1908, v. 28, 315–343.
81. Mason T. and Mason R. The journey to Palomar, PBS, DVD released on November 18, 2008 (90 minutes).
82. Newall H. F. George Ellery Hale. 1868-1938. *Obituary Not. Fell. Roy. Soc.*, 1939, v. 2(7), 522–526.
83. Gamow G. The birth and death of the sun: A lucid explanation of stellar evolution and atomic energy. New American Library, New York, N.Y., 1952.
84. Struve O. Stellar evolution: An exploration from the laboratory. Princeton University Press, Princeton, N.J., 1950.
85. Meadows A.J. Stellar evolution (2nd Edition), Pergamon Press, Oxford, U.K., 1972.

86. Arnett D. Supernovae and nucleosynthesis: An investigation of the history of matter, from the Big Bang to the present. Princeton University Press, Princeton, N.J., 1996.

87. Pagel B. E. J. Nucleosythesis and the chemical evolution of galaxies (2nd Edition), Cambridge University Press, Cambridge, U.K., 2009.

88. Gamow G. Nuclear energy sources and stellar evolution. *Phys. Rev.*, 1938, v. 53, 595–604.

89. Gamow G. Nuclear reactions in stellar evolution. *Nature*, 1939, v. 144, 620–622.

90. Bethe H.A. and Critchfield C.L. The Formation of Deuterons by Proton Combination. *Phys. Rev.*, 1938, v. 54, no. 4, 248–254.

91. Bethe H. A. Energy Production in Stars. *Phys. Rev.*, 1939, v. 55, no. 1, 103.

92. Bethe H.A. Energy Production in Stars. *Phys. Rev.*, 1939, v. 55, no. 5, 434–456.

93. von Weizsäcker C. F. Über Elementumwandlungen in Innern der Sterne II. *Physikalische Zeitschrift*, 1938, v. 39, 633–646.

94. Hoyle F. The synthesis of the elements from hydrogen. *Mon. Not. Roy. Astron. Soc.*, 1946, v. 106, 343–383.

95. Hoyle F. On nuclear reactions occurring in very hot stars. I. The synthesis of elements from carbon to nickel. *Astrophys. J. Suppl. Ser.*, 1954, v. 1, 121–146.

96. Burbidge M., Burbidge G.R., Fowler W.A. and Hoyle F. Synthesis of the elements in stars. *Rev. Mod. Phys.*, 1957, v. 29, no. 4, 547–654.

97. Wallerstein G., Iben I., Parker P., Boesgaard A.M., Hale G.M., Champagne A.E., Barnes C.A., Käppeler F., Smith V.V., Hoffman R.D., Timmes F.X., Sneden C., Boyd R.N., Meyer B.S. and Lambert D.L. Synthesis of the elements in stars: Forty years of progress. *Rev. Mod. Phys.*, 1997, v. 9, no. 4, 995–1084.

98. Bahcall J.N. Neutrinos from the Sun. *Sci. Am.*, 1969, v. 221, no. 1, 28–37.

99. Bahcall J.N. How the Sun shines.
www.nobelprize.org/nobel_prizes/physics/articles/fusion/index.html?print=1

100. Payne C.H. The relative abundances of the elements. Stellar Atmospheres. Harvard Observatory Monograph no. 1 (Harlow Shapley, Editor), Harvard University Press, Cambridge, MA, 1925 (reprinted in part in Lang K.R. and Gingerich O. A source book in astronomy and astrophysics, 1900–1975, Harvard University Press, Cambridge, MA, 1979, p. 245–248).

Structure Shape Evolution in Lanthanide and Actinide Nuclei

Khalaf A.M.* and Ismail A.M.†

*Physics Department, Faculty of Science, Al-Azhar University, Cairo, Egypt. E-mail: ali_khalaf43@hotmail.com
†Hot Laboratories Center, Atomic Energy Authority, Egypt, P.No. 13759, Cairo, Egypt. E-mail: ahmedismailph@yahoo.com

To give the characteristics of the evolution of the collectivity in even-even nuclei, we studied the behavior of the energy ratios R(4/2) and R(6/4). All chains of lanthanides begins as vibrational with R(4/2) near 2.0 and move towards rotational (R(4/2) $\longrightarrow$ 3.33) as neutron number increases. A rabid jump in R(4/2) near N=90 was seen. The plot of R(4/2) against Z shows not only the existence of a shape transitions but also the change in curvature in the data for N=88 and 90, concave to convex. For intermediate structure the slopes in E-GOS (E_γ over spin) plots range between the vibrator and rotor extremes. The abnormal behavior of the two-neutron separation energies of our lanthanide nuclei as a function of neutron number around neutron number 90 is calculated. Nonlinear behavior is observed which indicate that shape phase transition is occurred in this region. The calculated reduced B(E2) transition probabilities of the low states of the ground state band in the nuclei ^{150}Nd/^{152}Sm/^{154}Gd/^{156}Dy are analyzed and compared to the prediction of vibrational U(5) and rotational SU(3) limits of interacting boson model calculations.

1 Introduction

The interacting boson model (IBM) [1, 2] and the geometric collective model (GCM) [3–5] represent two major phenomenological approaches that successfully describe nuclear collectivity. While the IBM model is purely algebraic, based on a bosonized form of the many-body problem with even numbers of fermions, the GCM model follows from a geometric description of nuclei using the Bohr-Mottelson (BM) Hamiltonian [6].

Quantum phase transitions are of great interest in many areas of physics, and their manifestations vary significantly in different systems. For nuclear systems, the IBM reveals rich features of their shape phase transitions [7–16]. Three dynamical symmetries in the IBM were shown to correspond to three typical shape phase of nuclei, known as the spherical U(5) symmetry, axially deformed SU(3) symmetry and γ-soft deformed O(6) symmetry shapes. It is also known that phase transitions coincide with transitions between dynamical symmetries, with a first order phase transition taking place in the U(5)-SU(3) transition, and a second order phase transition happening in the U(5)-O(6).

A new class of symmetries that applies to systems localized at the critical points was proposed. In particular the critical symmetry E(5) [17] has been suggested to describe critical points in the phase transition from spherical vibrator U(5) to γ-unstable rotor O(6) shapes, while X(5) [18] is designed to describe systems lying at the critical point in the transition from spherical to axially deformed systems. These are based originally on particular solutions of the Bohr-Mottelson differential equations, but are usually applied in the context of the IBM [1], since the IBM provides a simple but detailed framework in which first and second order phase transitions can be studied. In the IBM language, the symmetry E(5) corresponds to the critical point between U(5) and O(6) symmetry limits, while X(5) symmetry should describe the phase transition region between the U(5) and the SU(3) dynamical symmetries.

The purpose of this paper is to disuse the main concepts of the rapid changes in structure of lanthanide and actinide nuclei by using some good indicators like energy ratios, two-neutron separation energies and reduced electric quadruple transition probabilities.

2 Energy Ratios and Nuclear Shape Transition

Nuclear shape phases are the manifestation of the collective motion modes of nuclei. One of the best signatures of shape transition is the behavior of the ratio between the energies of the first 4^+ and 2^+ states

$$R(4/2) = \frac{E(4_1^+)}{E(2_1^+)} \tag{1}$$

along the isotopic chain. The members of vibrational nuclei have excitation energies

$$E(I) = C(I), \tag{2}$$

where C is the vibrational constant. So that the energy ratios are

$$R((I+2)/I)_{vib} = \frac{I+2}{I}. \tag{3}$$

The yrast energies of the harmonic vibrator can be written as

$$E(I) = nE(2_1^+), \tag{4}$$

where n is the phonon number. The γ-ray energies within the yrast band are given by

$$\begin{aligned} E_\gamma(I) &= E(I) - E(I-2) \\ &= E(2_1^+). \end{aligned} \tag{5}$$

It is interesting to discuss the energy levels by plotting the ratio of E_γ(I) to spin I (E-Gamma Over Spin) (E-GOS) [19] against spin I. This is not helpful to identify the structure of the nucleus, but also to discern changes as a function of spin. Therefore, the E-GOS for vibrational nuclei can be written as

$$(E_\gamma/I)_{vib} = E(2_1^+)/I \tag{6}$$

which decreases hyperbolically from $E(2_1^+)/2$ to zero. In the rigid rotor, the energies of the yrast states are:

$$E(I) = AI(I+1), \tag{7}$$

where A is the rotational parameter ($A = \hbar^2/2J$, where J represents the moment of inertia), so that the energy ratios are

$$R((I+2/I))_{rot} = \frac{(I+2)(I+3)}{I(I+1)}. \tag{8}$$

Then The γ-ray energies within the yrast band are given by

$$E_\gamma(I) = A(4I-2) \tag{9}$$

and so the E-GOS is

$$\begin{aligned}(E_\gamma/I)_{rot} &= A\left(4-\frac{2}{I}\right)\\ &= \frac{E(2_1^+)}{3}\left(2-\frac{1}{I}\right).\end{aligned} \tag{10}$$

In units of A, this evolves from 3 for I=2 up to 4 for high I, and so gradually increasing and asymptotic function of I. Also E-GOS for γ-unstable nuclei is given by

$$(E_\gamma/I)_{\gamma-soft} = \frac{E(2_1^+)}{4}\left(1+\frac{2}{I}\right). \tag{11}$$

The R(4/2) varies from the value which correspond to vibrations around a spherical shape of vibrational nuclide R(4/2)=2 to the characteristic value for excitations of well-deformed rotor R(4/2)=3.33. That is, the energy ratio R(4/2) exhibits sharp change in rapid transitional region. Even-even nuclei can be classified roughly according to ratios R(4/2) as:

$1.0 < R(4/2) < 2.0$	for magic nuclei,
$2.0 < R(4/2) < 2.4$	for vibrational nuclei,
$2.4 < R(4/2) < 2.7$	for γ-unstable nuclei,
$2.7 < R(4/2) < 3.0$	for transitional nuclei,
$3.00 < R(4/2) < 3.33$	for rotational nuclei.

To give the characteristics of the evolution of the collectivity in even-even nuclei, we study the behavior of the energy ratios R(4/2) and R(6/4). For the nuclei included in our study, all chains of lanthanides begins as vibrational with R(4/2) near 2.0 and move towards rotational (R(4/2)$\longrightarrow$3.33) as neutron number increases. For intermediate structure the slopes in E-GOS plots range between the vibrator and rotor extremes. One particular case of interest is R(4/2)=3.0 which traditionally marks the boundary where axial rotation begins to set in. A very general phenomenological model is that of the an harmonic vibrator (AHV) [20]. In this model the yrast energies are given by

$$E(I=2n) = nE(2_1^+)\frac{n(n-1)}{2}\epsilon_4, \tag{12}$$

where

$$\epsilon_4 = E(4_1^+) - 2E(2_1^+) \tag{13}$$

is the an harmonically of the 4^+ state, that is, its deviation in energy from twice the 2^+ energy, and n = I/2, n is the phonon number in a vibrational nucleus. For $\epsilon_4 = 0$ equation (12) gives the harmonic vibrator

$$E(I) = \frac{1}{2}E(2_1^+)I \quad (R(4/2) = 2). \tag{14}$$

For $\epsilon_4 = (4/3)E(2_1^+)$, it gives the rigid rotor expression

$$E(I) = \frac{1}{6}E(2_1^+)I(I+1) \quad (R(4/2) = 10/3). \tag{15}$$

For $\epsilon_4 = E(2_1^+)$, it gives

$$E(I) = \frac{1}{8}E(2_1^+)I(I+2) \quad (R(4/2) = 3.0). \tag{16}$$

E(I)/I is constant and that the E-GOS plots is flat. So, interestingly the phase transition point (R(4/2) 3.0) roughly serves to section E-GOS plots into two classes of increasing and decreasing with I, so that nuclei on the vibrator side of the phase transition are down-sloping while these to the rotor side are up-sloping.

The systematics of energy ratios of successive levels of collective bands in medium and heavy mass even-even nuclei were studied [21]. A measure of their deviation from the vibrational and rotational limiting value was found to have different magnitude and spin dependence in vibrational, rotational and γ-unstable nuclei. For a given band for each spin I, the following ratios were constructed to define the symmetry for the excited band of even-even nuclei

$$\begin{aligned}r((I+2)/I) &= \frac{R((I+2)/I)_{exp} - R((I+2)/I)_{vib}}{R((I+2)/I)_{rot} - R((I+2)/I)_{vib}}\\ &= \frac{R((I+2)/I)_{exp} - (I+2)/I}{\dfrac{2(I+2)}{I(I+1)}},\end{aligned} \tag{17}$$

where $R((I+2)/I)_{exp}$ is the experimental value of the ratio. In equation (17), the value of energy ratios, r have changed between 0.1 and 1 for yrast bands of even-even nuclei. The ratio r should be close to one for a rotational nucleus and close to zero for a vibrational nucleus, while it should have intermediate values for γ-unstable nuclei:

$0.10 \le r \le 0.35$	for vibrational nuclei,
$0.4 \le r \le 0.6$	for transitional nuclei,
$0.6 \le r \le 1.0$	for rotational nuclei.

3 Electromagnetic Transition Strengths

When the nucleus is deformed it acquires an electric-multiple moment. Consequently as it oscillates, in $\lambda\mu$ mode, it emits electric $\lambda\mu$ radiation. Now to calculate the radiative transition rates between vibrational states, we need the nuclear electric multiple operator $\hat{\mathcal{M}}$. This is given by

$$\hat{\mathcal{M}}(E\lambda,\mu) = \int_{\tau} d\tau \rho_c(r) r^{\lambda} Y_{\lambda\mu}(\theta,\phi), \tag{18}$$

$\rho_c(r)$ is the charge density of the nucleus. The electric multipole moment is defined by $\hat{Q}_\lambda$

$$\hat{Q}_\lambda = \left(\frac{16\pi}{2\lambda+1}\right)^{1/2} \mathcal{M}(E\lambda,0). \tag{19}$$

We now discuss the electric quadruple moment ($\lambda = 2$) in more detail because the electric quadruple moment Q_2 of a nucleus is a measure of the deviation of the charge distribution from spherical symmetry. We define the reduced transition probability as:

$$\begin{aligned} B(E2, I_i \longrightarrow I_f) &= \textstyle\sum_{Mf} |\langle I_i M_i | Q_2 | I_f M_f\rangle|^2 \\ &= \frac{1}{2I_i+1} |\langle I_i \| Q_2 \| I_f\rangle|^2, \end{aligned} \tag{20}$$

where $|\langle I_i \| Q_2 \| I_f\rangle|$ is a reduced matrix element defined by the Wigner-Eckart theorem

$$\langle I_i M_i | \mathcal{M}(E\lambda,\mu) | I_f M_f\rangle| = \langle I_i M_i \lambda\mu | I_f M_f\rangle \frac{|\langle I_i \| \mathcal{M}(E\lambda) \| I_f\rangle|}{(2I_i+1)^{1/2}}.$$

The reduced transition probability $B(E2, I_i K \longrightarrow I_f K)$ for an electric quadruple transition between two members of same rotational band with quantum number K is:

$$B(E2, I_i K \longrightarrow I_f K) = \frac{5}{16\pi} e^2 Q_0^2 \langle I_i K 20 | I_f K\rangle^2, \tag{21}$$

where Q_0 is the transition intrinsic quadruple moment and we have used

$$\sum_{m_1 m_2, m} |\langle I_1 m_1 I_2 m_2 | I m\rangle|^2 = 2I + 1. \tag{22}$$

For even-even nuclei, $K = 0$ and when $I_i = I$ and $I_f = I - 2$, we get the familiar relations between $B(E2, I \longrightarrow I-2)$ and the intrinsic quadruple moment Q_0 are:

$$B(E2, I \longrightarrow I-2) = \frac{5}{16\pi} e^2 Q_0^2 \frac{3}{2} \frac{I(I-1)}{2(2I-1)(2I+1)}. \tag{23}$$

As a special case for the transition $2^+ \longrightarrow 0^+$, yields

$$B(E2, 2^+ \longrightarrow 0^+) = \frac{5}{16\pi} e^2 Q_0^2. \tag{24}$$

For the transition $I_i = I$ and $I_f = I + 2$, yields

$$B(E2, I \longrightarrow I+2) = \frac{5}{16\pi} e^2 Q_0^2 \frac{3}{2} \frac{(I+2)(I+1)}{2(2I+1)(2I+2)} \tag{25}$$

and for special case for the transition $0^+ \longrightarrow 2^+$, yields

$$B(E2, 0^+ \longrightarrow 2^+) = \frac{5}{16\pi} e^2 Q_0^2. \tag{26}$$

That is

$$B(E2, 2^+ \longrightarrow 0^+) = 0.2\ B(E2, 0^+ \longrightarrow 2^+). \tag{27}$$

From equation (21), the intrinsic quadruple moment Q_0 for a $K = 0$ band of an axially symmetric rotor is extracted. For the special transition $0^+ \longrightarrow 2^+$, we get

$$eQ_0 = \left[\frac{16\pi}{5} B(E2, 0^+ \longrightarrow 2^+)\right]^{1/2} \tag{28}$$

in units of 10^{-24} cm^2.

The electric reduced transition probability $B(E\lambda)$ can be obtained from the transition probability per unit time for emission of photon of energy $\hbar\omega$, angular momentum λ and of electric type with the nucleus going from a state i to a state f defined by

$$T(E\lambda) = \frac{8\pi(\lambda+1)}{\lambda[(2\lambda+1)!!]^2} \frac{1}{\hbar} \left(\frac{E_\gamma}{\hbar c}\right)^{(2\lambda+1)}. \tag{29}$$

$T(E\lambda)$ for electric quadruple has the from

$$T(E2) = \frac{4\pi}{75} \frac{1}{\hbar} \left(\frac{E_\gamma}{\hbar c}\right)^5 B(E2). \tag{30}$$

For the quadruple transition $T(E2)$ can be derived experimentally from the relation

$$T(E2, 2^+ \longrightarrow 0^+) = \frac{ln2}{(1+\alpha)\tau_{1/2}}, \tag{31}$$

where α is the total conversion coefficient taken from the tabulated values given by Rose [22] and $\tau_{1/2}$ is the half life time. From equations (30) and (31), one can find $B(E2)$:

$$\begin{aligned} B(E2, 2^+ \longrightarrow 0^+) &= \frac{75\hbar}{4\pi} \left(\frac{\hbar c}{E_{2^+}}\right)^5 \frac{ln2}{(1+\alpha)\tau_{1/2}} \\ &= 0.565502 \left(\frac{100}{E_{2^+}}\right)^5 \frac{1}{(1+\alpha)\tau_{1/2}}, \end{aligned} \tag{32}$$

where $B(E2)$ is in units of e^2b^2 when E_{2^+} is in units of MeV and $\tau_{1/2}$ in units of nanosecond.

4 The two–neutron Separation Energies

The energy required to remove a neutron from a nucleus with Z proton and N neutron is called separation energy and is defined as:

$$S_n(Z,N) = [M(Z,N-1) + M_n - M(Z,N)]C^2. \tag{33}$$

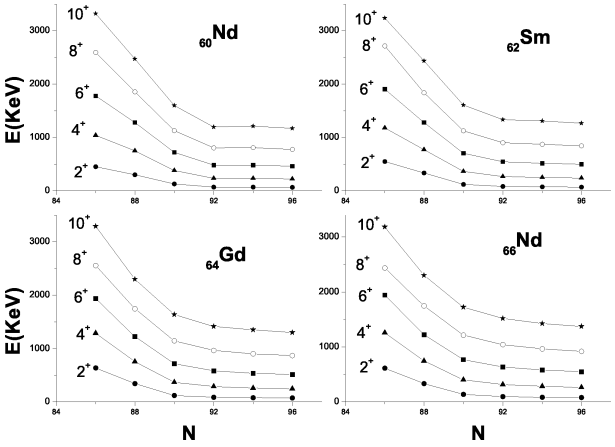

Fig. 1: Systematics of low-lying yrast level energies in even-even lanthanides Nd/Sm/Gd/Dy isotopes. The $2^+, 4^+, \ldots, 10^+$ level energies are plotted. The states are labeled by I^π.

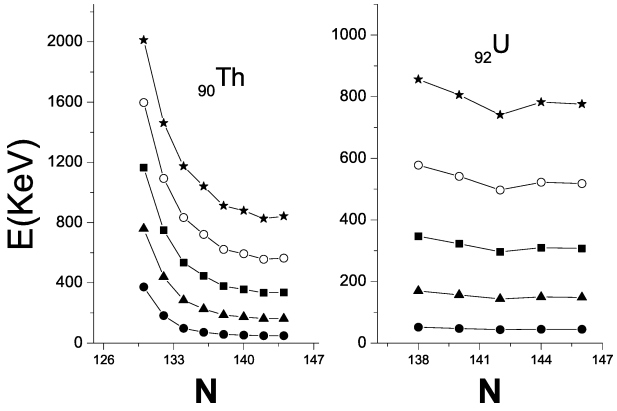

Fig. 2: The same as Fig. (1) but for actinides Th/U isotopes.

This expression can be rewritten in the form of binding energy as:

$$S_n(Z,N) = B(Z,N) - B(Z,N-1). \quad (34)$$

The definition of the two-neutron separation energies is the following:

$$S_{2n} = B(N) - B(N-1), \quad (35)$$

where N denotes the number of valence nucleon pairs and it is assumed that we are treating nuclei belonging to the first half of the neutron shell (50 - 82) filling up with increasing mass number.

5 Numerical Calculations and Discussions

The systematics of the excitation energies of the low-lying states as a function of neutron number changing from 84 to 100 in the even-even lanthanides Nd/Sm/Gd/Dy isotopes in the mass region 144–166 and the actinide Th/U isotopes in the mass region 224–238 are presented in Figures (1,2). Only the yrast state of positive parity and spin $I^\pi = 2^+, 4^+, 6^+, 8^+$ and 10^+ has been included.

The trend of increasing excitation energy of 2^+ state with decreasing neutron number, implying a corresponding fall in deformation as the N = 82 shell closure is approached. The energies of the 4^+ and 6^+ states also display the same trend. For lanthanides isotopes we can see that the energy values for each spin I states change almost linearly for $N \leq 88$ and become quite flat for $N \geq 90$. This is consistent with the onset of the Z = 64 sub-shell effect. Furthermore, the linear falling of the energy value for each I state as N goes from 86 to 88 seems to justify the linear variation of the effective proton-boson number in each isotope series.

As an example Figure (1) shows that the limits (spherical shape and well deformed rotor) are fulfilled in the Neodymium ^{144}Nd and $^{152-156}$Nd isotopes respectively, and also that there is a smooth transition between them. The ^{148}Nd isotopes could be considered as a transitional nucleus in the calculations. A rapid rise in R(4/2) between N = 88 and 90 is shown, where it increases from values of ≃2.3 typical of actual vibrational nuclei to 3.0, the traditional borderline value separating spherical from deformed nuclei to ≃3.3 the limiting value of the axial rotor model. As a matter of fact, if we compare the X(5) results (first order phase transition from a spherical vibrator to an axially deformed rotor is called X(5)) with the energy levels in ^{148}Nd, we find striking similarities, it suggested that the nucleus ^{148}Nd display the X(5) symmetry.

The nature of the low-lying states in our lanthanides and actinides chains of isotopes can be illustrated in Figures (3,4) by examining the ratios of the excitation energies R(4/2) and R(6/4) as a function of neutron number. The limiting values for R(4/2) and R(6/4) for harmonic vibrator are 2.0 and 1.5 and for rigid symmetric rotor are 0.33 and 2.1 respectively.

In lanthanides the calculated values increases gradually from vibrational value to transitional value near N=90 to rotor

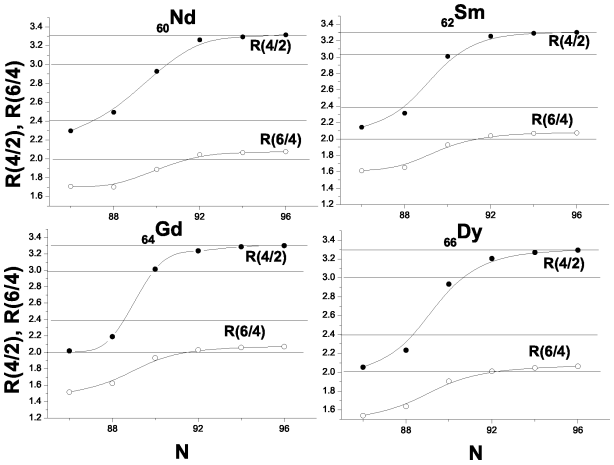

Fig. 3: Evolution of energy ratios R(4/2) and R(6/4) for lanthanides Nd/Sm/Gd/Dy isotopes as function of increasing neutron number.

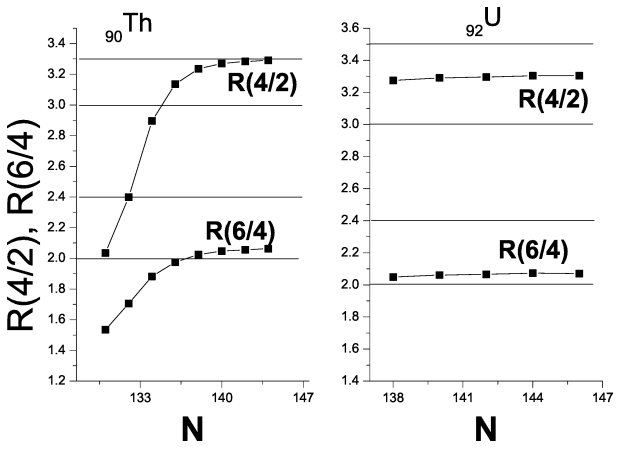

Fig. 4: The same as Fig. (3) but for actinides Th/U isotopes.

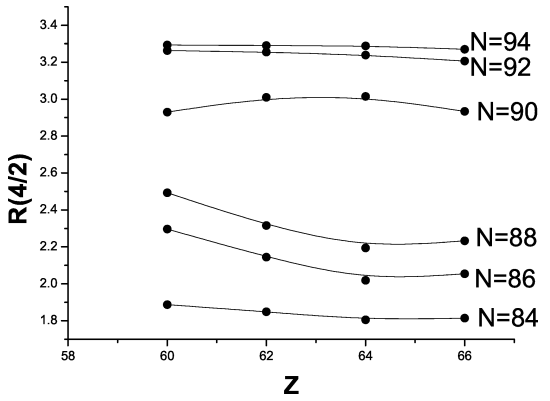

Fig. 5: The plot of R(4/2) values in the Nd/Sm/Gd/Dy region against Z. We see change in curvature in the data for N=88 and N=90 concave to convex.

value in the heavier isotopes. The energy ratios R(4/2) and R(6/4) for even A, N=88 isotopes are essentially constant for Sm, Gd and Dy.

The same data for lanthanides is plotted between R(4/2) against Z instead of N, see Figure (5). We see a rapid jump in R(4/2) near N=90. Here, the plot of R(4/2) against Z shows not only the existence of a shape transitions but also the change in curvature in the data for N=88 and 90, concave to convex. For Gd nuclei for N≤88 the behavior is typically closed shell, while for $N \geq 90$ the behavior appears to be near mid shell.

The nuclei of lanthanides region would therefore be candidates for a shape transition from vibrator to axially rotator and the N = 90 isotopes ^{150}Nd, ^{152}Sm, ^{154}Gd and ^{156}Dy are ideal candidates for X(5). Historically, sensitive studies [23] of the ^{152}Sm level scheme led to a suggestion that this nucleus gave evidence for a first order phase transition [24], its R(4/2) value is intermediate between vibrator and rotor [25]. Additional X(5) candidate in the lanthanides region have subsequently been identified in ^{150}Nd [26], ^{154}Gd [27], ^{156}Dy [28] and ^{162}Yb [29]. Fig. (6) shows R(I/2) and E-GOS plots for a vibrator, a rotor and R(4/2)=3 modes.

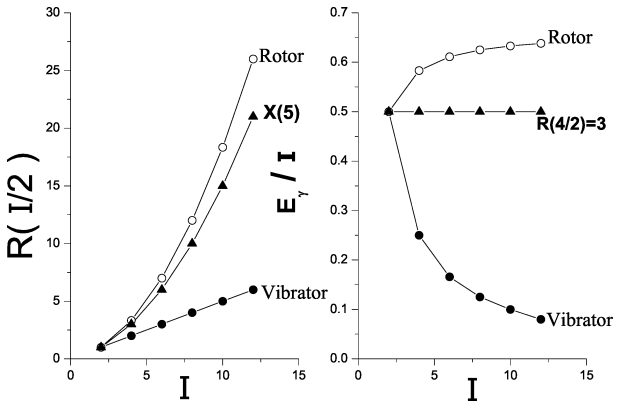

Fig. 6: Comparison of R(I/2) and E-GOS plots for three kinds of collective modes vibrator, rotor and R(4/2)=3 modes.

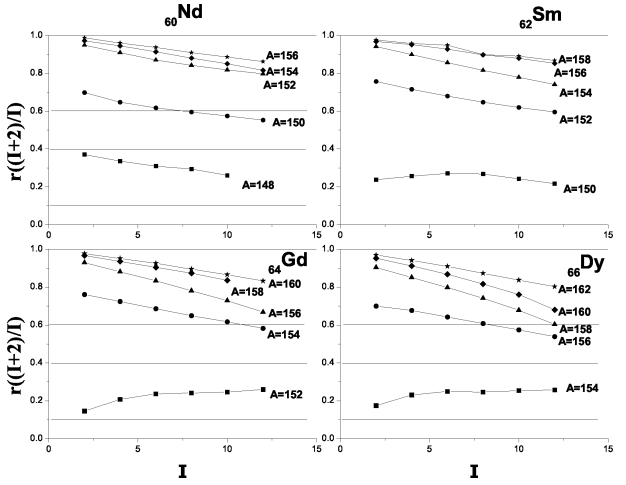

Fig. 7: The $r((I+2)/I)$ energy ratios for the ground state bands of even-even Lanthanides Nd/Sm/Gd/Dy isotopes as a function of spin I.

To investigate the dependence of energy ratios on the angular momentum, the useful criterion r((I + 2)/I) are examined for distinguishing between different kinds of collective behavior. In Figures (7,8) we show the results of our calculations for the ground state bands of the selected lanthanides and actinides isotopes. The study supports the interpretation of ^{150}Nd and ^{152}Sm as a critical point nucleus. Hence, the isotopes ^{150}Nd and ^{152}Sm are associated to X(5) symmetry. For the vibrational nuclei ^{152}Gd and ^{154}Dy, the ratios r((I + 2)/I) start with a small value and then increases with I, more rapidly in the beginning and slower at higher I's. On the other hand for rotational nucleus ^{162}Dy the ratios r((I + 2)/I) start with a value very close to one and then constantly decrease.

As an example, the abnormal behavior of the two-neutron

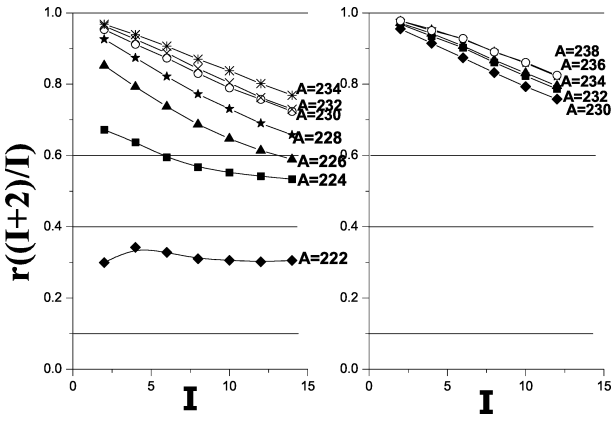

Fig. 8: The same as Fig. (7) but for Actinides Th/U isotopes.

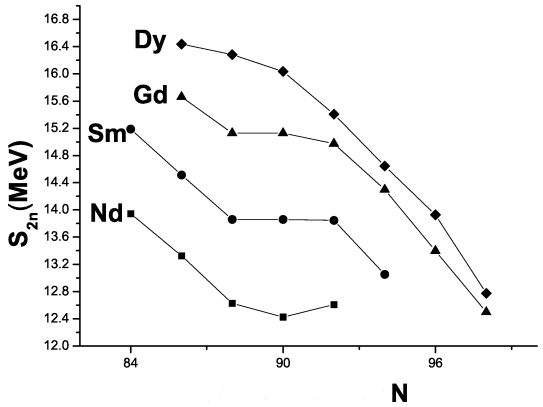

Fig. 9: Two-neutron separation energies S_{2n} for the chains Nd/Sm/Gd/Dy isotopes as a function of the number of neutrons.

separation energies S_{2n} of nuclei Nd/Sm/Gd/Dy as a function of neutron number around neutron number 90 is illustrated in Fig. (9), the nonlinear behavior of S_{2n} indicates that shape phase transition may occur in this region. It is commonly assumed that the ratio of the B(E2) reduced transition probabilities between the levels of the ground state band takes the values between vibrational and rotational limits. In the interacting boson model IBM [1] both these limits are corrected because the number of the quadruple bosons cannot exceed some maximum value N.

In the U(5) vibrational limit of IBM,

$$\frac{B(E2, I+2 \to I)}{B(E2, 2^+ \to 0^+)} = \frac{1}{2}(I+2)\left(1-\frac{1}{2N}\right)$$

and in the SU(3) rotational limit of IBM,

$$\frac{B(E2, I+2 \to I)}{B(E2, 2^+ \to 0^+)} = \frac{15}{2}\left(1-\frac{1}{2N}\right)\left(1-\frac{1}{2N+3}\right)\frac{(I+2)(I+1)}{(2I+3)(2I+5)}.$$

Our GCM calculated values of these ratios are put between these limits, *i.e.*, the IBM calculations can reproduce the E2 transition probabilities.

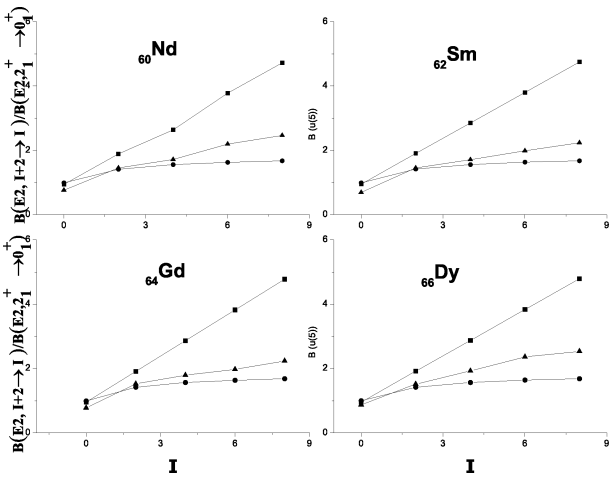

Fig. 10: The ratio $\frac{dB(E2,I+2\longrightarrow I)}{B(E2,2^+\longrightarrow 0^+}$ of reduced transition probabilities between the levels of the ground state band of ^{150}Nd, ^{152}Sm, ^{154}Gd and ^{156}Dy as compared to the U(5) and SU(3) of IBM calculations (• for U(5), ∘ for SU(3) and x for present calculation).

Table 1: The GCM parameters as derived in fitting procedure used in the calculation.

Nucleus	I	U(5) Vibrator	SU(3) Rotor	Present calculations
^{150}Nd (N=9)	0	0.94444	0.98941	0.75812
	2	1.88888	1.41345	1.45375
	4	2.63333	1.55677	1.71683
	6	3.77777	1.62962	2.19186
	8	4.72222	1.67381	2.46675
^{152}Sm (N=10)	0	0.95	0.99130	0.68900
	2	1.90	1.41614	1.45137
	4	2.85	1.55973	1.71262
	6	3.80	1.63272	1.98838
	8	4.75	1.67700	2.23512
^{154}Gd (N=11)	0	0.95454	0.99272	0.77300
	2	1.90909	1.41818	1.52393
	4	2.86363	1.56197	1.79560
	6	3.81818	1.63507	1.97412
	8	4.77272	1.67941	2.23803
^{156}Dy (N=12)	0	0.95833	0.99382	0.87381
	2	1.91666	1.41975	1.51345
	4	2.87500	1.56370	1.92725
	6	3.83333	1.63688	2.35673
	8	4.79166	1.68127	2.53512

The calculated $B(E2, I+2 \longrightarrow I)/B(E2, 2^+ \longrightarrow 0^+)$ ratios using GCM for the ground state bands of the low-lying state are presented in Table (1) and Fig. (10) together with

the results for the vibrator and rotor limits of IBM for ^{150}Nd, ^{152}Sm, ^{154}Gd and ^{156}Dy.

Submitted on January 15, 2013 / Accepted on January 21, 2013

References

1. Iachello F. and Arima A. The Interacting Boson Model. Cambridge University Press, Cambridge, England, 1987.
2. Frank A. and VanIsacker P. Algebraic Methods in Molecular and Nuclear Structure Physics. Wiley, New York, 1994.
3. Eisenberg J. and Greiner W. Nuclear Theory, Vol. I, Nuclear Models: Collective and Single-Particle Phenomena. North-Holland, Amsterdam, 1987.
4. Troltenier D., Hess P.O. and Maruhn J. Computational Nuclear Physics, Vol. I, Nuclear Structure. Springer, Berlin, Heidelberg, New York, 1991.
5. Troltenier D. Das Generalisierte Kollektivmodell. Frankfurt am Main, Germany, Report No. GSI-92-15, 1992.
6. Bohr A. and Mottelson. Nuclear Structure v. II. Benjamin, New York, 1975.
7. Jolie J. et al. Two-Level Interacting Boson Models Beyond The Mean Field. *Physical Reveiw*, 2007, v. C75, 014301R–014310R.
8. Iachello F. and Zamfir N.V. Quantum Phase Transitions in Mesoscopic Systems. *Physical Review Letters*, 2004, v. 92(3), 212501–212504.
9. Cejnar P., Heinze S. and Dobes J. Thermodynamic Analogy for Quantum Phase Transitions at Zero Temperature. *Physical Review*, 2005, v. C71, 011304R–011309R.
10. Rowe D.J. Quasi Dynamical Symmetry in an Interacting Boson Model Phase Transition. *Physical Review Letters*, 2004, v. 93, 122502–122505.
11. Liu Y.X., Mu L.Z. and Wei H. Approach to The Rotation Driven Vibrational to Axially Rotational Shape Phase Transition Along The Yrast Line of a Nucleus. *Physics Letters*, 2006, v. B633, 49–53.
12. Zhang Y., Hau Z. and Liu Y.X. Distinguishing a First Order From a Second Order Nuclear Shape Phase Transition in The Interacting Boson Model. *Physical Review*, 2007, v. C76, 011305R–011308R.
13. Arios J.M., Dukelsky J. and Garcia-Ramos J.E. Quantum Phase Transitions in the Interacting Boson Model: Integrability, Level Repulsion and Level Crossing. *Physical Review Letters*, 2003, v. 91, 162502–162504.
14. Garcia-Ramos J.E. et al. Two-Neutron Separation Energies, Binding Energies and Phase Transitions in The Interacting Boson Model. *Nuclear Physics*, 2001, v. A688, 735–754.
15. Liu M.L. Nuclear Shape-Phase Diagrams. *Physical Review*, 2007, v. C76, 054304–054307.
16. Heyde K. et al. Phase Transitions Versus Shape Coexistence. *Physical Review*, 2004, v. C69, 054304–054309.
17. Iachello F. Dynamic Symmetrcies at The Critical Point. *Physical Review Letters*, 2000, v. 85, 3580–3583.
18. Iachello F. Analytic Prescription of Critical Point Nuclei in a Spherical Axially Deformed Shape Phase Transtion. *Physical Review Letters*, 2001, v. 87, 052502–052506.
19. Regan P.H. et al. Signature for Vibrational to Rotational Evolution Along the Yrast Line. *Physical Review Letters*, 2003, v. 90, 152502–152505.
20. Zamfir N.V. et al. Study of Low-Spin States in ^{122}Cd. *Physical Review*, 1995, v. C51, 98–102.
21. Bonatsos D. and Skoures L.D. Successive Energy Ratios in Medium- and Heavy-Mass Nuclei as Indicators of Different Kinds of Collectivity. *Physical Review*, 1991, v. C43, 952R–956R.
22. Rose M.E. Internal Conversion Coefficients. Amsterdam, North Holland . North-Holland, Publishing Company 1958.
23. Casten R.F. The First Excited 0^+ State in ^{152}Sm. *Physical Review*, 1998, v. C57, 1553R–1557R.
24. Casten R.F. Kusnezov D. and Zamfir N.V. Phase Transitions in Finite Nuclei and The Integer Nucleon Number Problem. *Physical Review Letters*, 1999, v. 82, 5000–5003.
25. Clark R.M. et al. Searching For X(5) Behavior in Nuclei. *Physical Review*, 2003, v. C68, 037301–037304.
26. Krücken R. et al. B(E2) Values in ^{150}Nd and The Critical Point Symmetry X(5). *Physical Review Letters*, 2002, v. 88, 232501–232501.
27. Tonev D. et al. Transition Probabilities in ^{154}Gd: Evidence for X(5) Critical Point Symmetry. *Physical Review*, 2004, v. C69, 034334–034339.
28. Caprio M.A. et al. Low-Spin Structure of 156Dy Through γ-ray Spectroscopy. *Physical Review*, 2002, v. C66, 054310–054328.
29. McCutchan E.A. et al. Low Spin States in ^{162}Yb and The X(5) Critical Point Symmetry. *Physical Review*, 2004, v. C69, 024308–024317.

Double Surface and Fine Structure

Janez Špringer

Cankarjeva cesta 2, 9250 Gornja Radgona, Slovenia, EU. E-mail: janez_springer@t-2.net

Previously [1], one concluded that the atomic world should be elliptic and therefore the present universe which on the macro level looks like Euclidean is obviously to be heterogeneous. In this paper, one tries to solve the enigma proposing the double elliptic-hyperbolic surface. As a result of the effort, a new candidate for the exact inverse fine structure constant is given: $\alpha^{-1} = 137\left(2 - 1/\sqrt{1+\pi^2/137^2}\right) = 137.0360062543\ldots$.

1 Theoretical background

Let us consider our experience of the world is not what that world in reality is but rather how it is observed and measured. The distinction between to observe and to measure is made in this paper. The former means to count the units in the image, denoted as the average $\overline{x}$. The latter means to count the units in the inverse image, denoted as the average $\overline{x^{-1}}$. For the different values of x_i we have to deal with the next inequality:

$$\overline{x} \times \overline{x^{-1}} \neq 1. \tag{1}$$

Then the surface we live on is not, for instance, the Euclidean plane or the sphere very close to it [1], but could be, instead of it, the double elliptic-hyperbolic surface which is observed as the Euclidean plane. The average sphere is not proposed to be the triple elliptic-Euclidean-hyperbolic surface unless the Euclidean plane is not assigned to have its own identity. Let us propose that this leaves a footprint in the inverse fine structure constant α^{-1} which is in some way observed. Actually in the observation we count the number of the length units λ which are correlated with the inverse fine structure α^{-1}:

$$\alpha^{-1}_{observed} = \alpha^{-1}_{euclidean} = \frac{\alpha^{-1}_{elliptic} + \alpha^{-1}_{hyperbolic}}{2}. \tag{2}$$

And the measured elliptic fine structure constant on the atom level does not reflect exclusively the elliptic sphere, since it is the mirror of the hyperbolic sphere, too. Let us propose that this leaves a footprint in the fine structure constant α which is in some way measured. Actually in the measurement we count the number of the inverse length units $\lambda^{-1} = mv/h$ which are correlated with the fine structure α:

$$\alpha_{measured} = \frac{\alpha_{elliptic} + \alpha_{hyperbolic}}{2}. \tag{3}$$

Consequently the different inverse fine structure constants are explicitly expressed as

$$\alpha^{-1}_{measured} = \alpha^{-1}_{elliptic}\left(2 - \frac{\alpha^{-1}_{elliptic}}{\alpha^{-1}_{euclidean}}\right), \tag{4a}$$

$$\alpha^{-1}_{elliptic} = \alpha^{-1}_{euclidean} - \sqrt{\alpha^{-1}_{euclidean}\left(\alpha^{-1}_{euclidean} - \alpha^{-1}_{measured}\right)}, \tag{4b}$$

$$\alpha^{-1}_{hyperbolic} = \alpha^{-1}_{euclidean} + \sqrt{\alpha^{-1}_{euclidean}\left(\alpha^{-1}_{euclidean} - \alpha^{-1}_{measured}\right)}, \tag{4c}$$

$$\alpha^{-1}_{sphere} = \alpha^{-1}_{euclidean} \mp \sqrt{\alpha^{-1}_{euclidean}\left(\alpha^{-1}_{euclidean} - \alpha^{-1}_{measured}\right)}. \tag{4d}$$

It is easily seen that if the measured inverse fine structure constant equals the observed Euclidean one, the elliptic and hyperbolic inverse fine structure constant are identical and no average makes sense. Only in that case what is observed and measured is also real.

Let us also recall the value of the hypothetical Euclidean inverse fine structure constant [1]:

$$\alpha^{-1}_{euclidean} = \sqrt{\pi^2 + 137^2}. \tag{5}$$

2 The fine structure constant and the Hydrogen atom

The elliptic sphere of the radius of about 3679 Compton wavelengths of the electron was proposed in the Hydrogen atom previously [1], based on the assumption that only one type of the sphere is possible. If the elliptic and hyperbolic sphere coexists, the fine structure constant is a mirror of their average geometry, and what results is a different sphere picture. Without going into the details of how it looks like, some calculations can be made.

2.1 Calculation of the sphere paths

Taking into account the equation (5) and inserting in the equations (4b) and (4c), the CODATA 2012 recommended $\alpha^{-1} = 137.035999074$ for the $\alpha_{measured}$, the elliptic and hyperbolic path s in the Hydrogen atom are given in units of Compton wavelengths of the electron as:

$$\begin{aligned} s_{elliptic}(\alpha^{-1}_{elliptic}) &= 136.988254898\cdots < n = 137 \\ s_{hyperbolic}(\alpha^{-1}_{hyperbolic}) &= 137.083776540\cdots \end{aligned} \tag{6}$$

The path on the elliptic sphere being smaller than the translation component n is not plausible and leads one to the conclusion that the recommended empirical value of α^{-1} should be of a little greater size.

2.2 Calculation of the inverse fine structure constants

The translation component $n = 137$ Compton wavelengths of the electron equals the elliptic circular path s and the latter

expresses the elliptic inverse fine structure constant [1]

$$\alpha^{-1}_{elliptic} = 137, \text{ since:}$$
$$n = s = 137 \text{ Compton wavelengths of the electron.} \tag{7}$$

The theoretical inverse fine structure constant deduced from the average path on the double elliptic-hyperbolic surface is given with the equations (4a) and (5):

$$\begin{aligned} \alpha^{-1}_{theoretical} &= 137\left(2 - 1/\sqrt{1+\pi^2/137^2}\right) \\ &= 137.0360062543\cdots < \alpha^{-1}_{euclidean} \end{aligned} \tag{8}$$

The calculated constant is a little greater than the recommended CODATA 2012 α^{-1} but smaller than the hypothetical Euclidean one given by (5). The hyperbolic inverse fine structure is given by (4c):

$$\alpha^{-1}_{hyperbolic} = 137.0720314399\cdots \tag{9}$$

3 Conclusion

According to the proposed model, the electron in the Hydrogen atom moves on the elliptic-hyperbolic double surface, since the measured inverse fine structure constant is smaller than the hypothetical Euclidean one. And we live in the apparent Euclidean macro-world, since the observed inverse fine structure constant does not seem to differ from the hypothetical Euclidean one. The difference between what is observed on the macro level and what is measured in the atom world implies that neither what is observed nor what is measured is real. If the elliptic and hyperbolic sphere can coexist in the present world, a new candidate for the exact inverse fine structure constant is given by

$$\alpha^{-1}_{theoretical} = 137\left(2 - 1/\sqrt{1+\pi^2/137^2}\right) = 137.0360062543\cdots$$

Dedication

This fragment is dedicated to my granddaughters Urša and Špela.

Submitted on January 28, 2013 / Accepted on February 7, 2013

References

1. Špringer J. Fine Structure Constant as a Mirror of Sphere Geometry. *Progress in Physics*, 2013, v. 1, 12–14.

LETTERS TO PROGRESS IN PHYSICS

Commentary Relative to the Distribution of Gamma-Ray Flares on the Sun: Further Evidence for a Distinct Solar Surface

Pierre-Marie Robitaille

Department of Radiology, The Ohio State University, 395 W. 12th Ave, Columbus, Ohio 43210, USA.
robitaille.1@osu.edu

High energy gamma-ray flares are almost always observed near the limb of the Sun and are seldom, if ever, visualized in the central region of the solar disc. As such, they exhibit a powerful anisotropy best explained by invoking a true photospheric surface. In this regard, the anisotropic nature of the gamma-ray emissions from high-energy flares constitute the eighteenth line of evidence that the Sun is condensed matter.

Every body has a surface.

St. Thomas Aquinas [1]

In the middle ages, as St. Thomas Aquinas was reflecting upon *The Infinity of God*, he was confronted with this objection relative to objects and their surfaces [1]. Thomas would answer that: "*It is one thing to be infinite in essence, and another to be infinite in magnitude*" [1]. Though nearly a millennium has passed since the Dominican Friar contemplated *The Infinity of God*, the fact remains that, in the physical world, one is primarily considering magnitude, not essence: on a macroscopic scale, every physical body does indeed have a surface. Failure to meet this criterion results in an assembly of many bodies.

These ideas have consequences for astronomy. Within the context of accepted solar models, the Sun must be viewed as an assembly of bodies, since it has long ago been deprived of a real surface by gaseous constructs [2].

Conversely, the author has argued that the Sun does indeed possess a real surface [3] and he has recently assembled a wide variety of proofs that highlight its condensed state of matter (see e.g. [4] and references therein). In this brief work, an 18^{th} line of evidence is provided.

In 1989, Erich Rieger published a paper in *Solar Physics* entitled "*Solar Flares: High Energy Radiation and Particles*" [5]. In this report, Rieger provided strong evidence that flares with emissions >10 MeV are visible only near the solar limb (see Fig. 1). Rieger's findings would be highlighted by R. Ramaty and G. M. Simnett in their review on accelerated particles in solar flares: "*Gamma-ray emitting flares are observed from sites located predominantly near the limb of the Sun (see, e.g. Rieger 1989). This effect was observed for flares detected at energies >0.3 MeV, but it is at energies >10 MeV that the effect is particularly pronounced ... Since in both of these cases the bulk of the emission is bremsstrahlung from primary electrons, these results imply that the radiating electrons are anisotropic*" [6, p. 237]. It was then postulated that: "*... the anisotropy could result from the mirroring of the charged particles in the convergent chromospheric magnetic fields*" [6, p. 237] based on a theoretical analysis by Miller and Ramaty [7]. These authors comment that the emissions are "*... strongly anisotropic, with more emission in the directions tangential to the photosphere than in directions away from the Sun*" [7]. In order to account for the anisotropy of the gamma-ray emission from high energy solar flares, they invoke electron transport in the coronal region and magnetic mirroring of converging magnetic flux tubes beneath the transition region [7]. As the gaseous models of the Sun cannot support the existence of a real surface, then another mechanism must be created to "act as a surface".

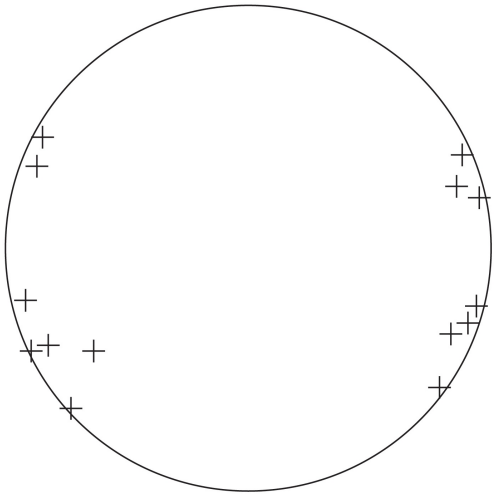

Fig. 1: Schematic representation of the relative position of flares with >10 MeV of energy on the solar disk displaying their predominance near the limb. This figure is meant only for illustrative purposes and is an adaptation based on Fig. 9 in [5] which should be examined for exact flare locations.

Within the gaseous models, the photosphere merely represents a region of increasing opacity, best regarded as an "optical illusion" [3]. The gaseous Sun possesses no sudden change in density which could allow tangential emission to its surface. In fact, modern solar models assume a density of only 10^{-7} g/cm^3 for the photosphere [8, p. 32], a density

lower than some of our earthly vacuums. Hence the use of magnetic mirroring and the convergence of field lines in order to generate surface effects in the absence of condensed matter.

In the end, the simplest way to account for the strongly anisotropic nature of high energy solar flares is to recognize the existence of a discrete surface on the Sun. This most elegantly explains why the emissions are *tangential to the photosphere*. As flares rise from the solar interior [4] they eventually arrive at the photospheric layer. High energy gamma rays are emitted tangentially to this boundary, as a real physical surface, not to an illusion [3], has been encountered.

Acknowledgment

Luc Robitaille is recognized for the preparation of Figure 1.

Dedication

Dedicated to Dominican Friars of the Province of St. Joseph.

Submitted on: January 17, 2013 / Accepted on: January 18, 2013
First published in online on: January 28, 2013

References

1. Aquinas, Thomas. ST I,7,iii: Cosimo Inc., New York, 2007.
2. Robitaille P.M. A thermodynamic history of the solar constitution — I: The journey to a gaseous Sun. *Progr. Phys.*, 2011, v. 3, 3–25.
3. Robitaille P.M. On the Presence of a Distinct Solar Surface: A Reply to Hervé Faye. *Progr. Phys.*, 2011, v. 3, 75–78.
4. Robitaille J.C. and Robitaille P.M. Liquid Metallic Hydrogen III: Intercalation and lattice exclusion versus gravitational settling and their consequences relative to internal structure, surface activity, and solar winds in the Sun. *Progr. Phys.*, 2013, v. 2, in press.
5. Rieger E. Solar flares: High-energy radiation and particles. *Solar Phys.*, 1989, v. 121, 323–345.
6. Ramata R. and Simnett G.M. Accelerated particles in solar flares. In: *The Sun in Time* (C.P. Sonett, M.S. Giampapa and M.S. Matthews, Eds.), The University of Arizona Press, Tuscon, AZ, 1991, p. 232–259.
7. Miller J.A. and Ramaty R. Relativistic electron transport and Bremsstrahlung production in solar flares. *Astrophys. J.*, 1989, v. 344, 973–990.
8. Bhatnagar A. Instrumentation and observational techniques in solar astronomy. In: *Lectures on Solar Physics* (H.M. Antia, A. Bhatnagar and R. Ulmschneider, Eds.), Springer, Berlin, 2003, p. 27–79.

LETTERS TO PROGRESS IN PHYSICS

Commentary Relative to the Seismic Structure of the Sun: Internal Rotation, Oblateness, and Solar Shape

Pierre-Marie Robitaille

Department of Radiology, The Ohio State University, 395 W. 12th Ave, Columbus, Ohio 43210, USA.
robitaille.1@osu.edu

Helioseismological studies have the ability to yield tremendous insight with respect to the internal structure and shape of the solar body. Such observations indicate that while the convection zone displays differential rotation, the core rotates as a rigid body. The latter is located below the tachocline layer, where powerful shear stresses are believed to occur. Beyond simple oblateness, seismological studies indicate that the Sun displays significant higher order shape terms (quadrupole, hexadecapole) which may, or may not, vary with the solar cycle. In this work, such seismological findings are briefly discussed with the intent of highlighting that 1) the differential rotation of the convection zone, 2) the rigid body rotation of the core, 3) the presence of the tachocline layer and 4) the appearance of higher order shape terms, all lend support to the idea that the solar body is composed of material in the condensed state. In this regard, the existence of the tachocline layer in the solar interior and the solid body rotation of the core constitute the nineteenth and twentieth lines of evidence that the Sun is condensed matter.

> *In brief, every rotating body conducts itself either as if it is it were purely liquid, or as if it were purely gaseous; there are no intermediate possibilities. Observational astronomy leaves no room for doubt that a great number of stars, perhaps even all stars . . . behave like liquids rather than gases.*
>
> Sir James Hopwood Jeans, 1929 [1]

For much of his life, James Jeans believed that stars were rotating liquids [1, 2]. On the basis of the tremendous abundance of binary systems [2], he had claimed that there could be no doubt of their condensed nature. Yet, in the paragraph which followed that quoted above, Jeans also argued: *"we are totally unable to check our theoretical results by observation"* [1, p. 219]. This apparent contradiction was previously highlighted by Alan B. Whiting [3, p. 209]. Eventually, Jeans lost sight of the observational evidence which had so convinced him. By 1944, he had abandoned liquid stars [2,4] and so did astrophysics; although in the 1960s, Subrahmanyan Chandrashekar would devote nine years of his life to the study of rotating liquid bodies [4,5]. With time however, astronomy would add to the arsenal of evidence that the Sun was liquid (see [6–8] and references therein).

Seismology, the study of low frequency waves within condensed matter, would also contribute to our understanding [9, 10]. Indeed, the mere application of seismology to the Sun has been heralded as a proof for condensed matter (see proof 5 in [8]). It is not reasonable to claim that the solar photosphere, with a density of only 10^{-7} g/cm^3 [11], can act as a mere optical illusion relative to the presence of a distinct surface [12], while at the same time forming the confines of a resonant cavity for seismological studies [13]. The author has already argued that it is not possible to conduct seismological observations on a surface whose density remains inferior to some of the best vacuums on Earth [8], despite the apparent agreement with the gaseous solar models [14, 15]. Seismology has been, and always will remain, linked to the study of condensed matter.

In this regard, seismology has brought some interesting insight into the internal structure of the Sun. The fact that the convection zone undergoes differential rotation appears well established, as is the presence of a prolate tachocline layer [9,10]. The tachocline region acts as a shear layer which separates the differential rotation in the convection zone from the solid body rotation observed in the solar core. Shear forces imply area and surface. As such, the presence of the tachocline layer in the solar interior is now advanced as the nineteenth line of evidence that the Sun is condensed matter. Furthermore, the solar core is rotating as a solid body (e.g. [10]) and this remains impossible for a gaseous object. Solid body rotation involves strong internal cohesive forces which gases cannot possess. Consequently, the solid body rotation of the solar core is now invoked as the twentieth line of evidence that the Sun is condensed matter.

Finally, it is well established that the Sun is not perfectly spherical but oblate (see [15, 16] and references therein). Indeed, the presence of solar oblateness could be related to Jean's arguments for liquid stars [2]. Since the creation of an oblate object requires internal cohesive forces which can only characterize a liquid or solid rotating sphere, solar oblateness has already been invoked as the eighth line of evidence that the Sun is condensed matter [8]. Yet, the solar shape is even more complex, characterized by quadrupolar and hex-

adecapolar terms [16], the latter of which appears dependent on the solar cycle. These additional features on the solar sphere served to complement the eighth line of evidence (solar shape [8]) that the Sun is condensed matter.

Dedication

This work is dedicated to Bernadine Healy†, Reed Fraley, Joan Patton, Bradford Stokes, and Kamilla Sigafoos for their administrative leadership in helping to create the world's first Ultra High Field MRI system at The Ohio State University.

Submitted on: January 25, 2013 / Accepted on: January 25, 2013
First published in online on: January 28, 2013

References

1. Jeans J. The universe around us. Cambridge University Press, 1929, p. 211.
2. Jeans J. Astronomy and Cosmogony. Cambridge University Press, 1928.
3. Whiting A.B. Hindsight and popular astronomy. World Scientific, New Jersey, 2011.
4. Robitaille P.M. A thermodynamic history of the solar constitution–II: The theory of the gaseous Sun and Jeans' failed liquid alternative. *Progr. Phys.*, 2011, v. 3, 3–25.
5. Chandrasekhar S. Ellipsoidal Figures of Equilibrium. Yale Unversity Press, New Haven, 1969.
6. Robitaille J.C. and Robitaille P.M. Liquid Metallic Hydrogen III: Intercalation and lattice exclusion versus gravitational settling and their consequences relative to internal structure, surface activity, and solar winds in the Sun. *Progr. Phys.*, 2013, v. 2, in press.
7. Robitaille P.M. Commentary relative to the distribution of gamma-ray flares on the Sun: Further evidence for a distinct solar surface. *Progr. Phys.*, 2013, v. 2, L1-L2.
8. Robitaille P.M. The solar photosphere: Evidence for condensed matter. *Progr. Phys.*, 2006, v. 2, 17–21 (also found in slightly modified form within *Research Disclosure*, 2006, v. 501, 31–34; title #501019).
9. Gough D.O. Seismology of the Sun and the distant stars. D. Reidel Publishing Company, Dordrecht, 1986.
10. Antia H.M. Solar interior and seismology. In: *Lectures on Solar Physics* (H.M. Antia, A. Bhatnagar and R. Ulmschneider, Eds.), Springer, Berlin, 2003, p. 80–126.
11. Bhatnagar A. Instrumentation and observational techniques in solar astronomy. In: *Lectures on Solar Physics* (H.M. Antia, A. Bhatnagar and R. Ulmschneider, Eds.), Springer, Berlin, 2003, p. 27–79.
12. Robitaille P.M. On the Presence of a Distinct Solar Surface: A Reply to Hervé Faye. *Progr. Phys.*, 2011, v. 3, 75–78.
13. Robitaille P.M. Stellar opacity: The Achilles heel of a gaseous Sun. *Progr. Phys.*, 2011, v. 3, 93–99.
14. Gough D.O., Kosovichev A.G., Toomre J., Anderson E., Antia H.M., Basu S., Chaboyer B., Chitre S.M., Christensen-Dalsgaard J., Dziembowski W.A., Eff-Darwich A., Elliott J.R., Giles P.M., Goode P.R., Guzik J.A., Harvey J.W., Hill F., Leibacher J.W., Monteiro M.J.P.F.G., Richard O., Sekii T., Shibahashi H., Takata M., Thompson M.J., Vauclair S., Vorontsov S.V. The seismic structure of the Sun. *Science*, v. 272, no. 5266, 1296–1300.
15. Godier S. and Rozelot J.P. The solar oblateness and its relationship with the structure of the tachocline and the Sun's subsurface. *Astron. Astrophys.*, 2000, v. 355, 365–374
16. Kuhn J.R., Bush R.I, Scheick X. and Scherrer P. The Sun's shape and brightness. *Nature*, 1998, v. 392, no. 6672, 155–157.

LETTERS TO PROGRESS IN PHYSICS

Commentary on the Radius of the Sun: Optical Illusion or Manifestation of a Real Surface?

Pierre-Marie Robitaille

Department of Radiology, The Ohio State University, 395 W. 12th Ave, Columbus, Ohio 43210, USA.
robitaille.1@osu.edu

In modern solar theory, the photospheric surface merely acts as an optical illusion. Gases cannot support the existence of such a boundary. Conversely, the liquid metallic hydrogen model supports the idea that the Sun has a distinct surface. Observational astronomy continues to report increasingly precise measures of solar radius and diameter. Even the smallest temporal variations in these parameters would have profound implications relative to modeling the Sun and understanding climate fluctuations on Earth. A review of the literature convincingly demonstrates that the solar body does indeed possess a measurable radius which provides, along with previous discussions (Robitaille P.M. On the Presence of a Distinct Solar Surface: A Reply to Hervé Faye. *Progr. Phys.*, 2011, v. 3, 75–78.), the twenty-first line of evidence that the Sun is comprised of condensed-matter.

> *But however difficult it may be for present theories to account for the tenuity of the solar atmosphere immediately above the photosphere, and however readily the same fact may be accounted for by the theory of Schmidt, it is certain that the observer who has studied the structure of the Sun's surface, and particularly the aspect of the spots and other markings as they approach the limb, must feel convinced that these forms actually occur at practically the same level, that is, that the photosphere is an actual and not an optical surface. Hence it is, no doubt, that the theory is apt to be more favorably regarded by mathematicians than by observers.*
>
> James Edward Keeler, 1895 [1]

James Edward Keeler was a distinguished observational astronomer [2]. Along with George Ellery Hale, he had established *The Astrophysical Journal* in 1895 [2]. In the first volume of this journal, Keeler objected to Schmidt's model of a fully gaseous Sun whose surface merely represented an optical illusion (see [3] for a full discussion). Hale echoed Keeler's objections stating, *"As a theoretical discussion the theory is interesting and valuable, but few observers of the Sun will consider it capable of accounting for the varying phenomena encountered in their investigations"* [4]. Thus, two of the greatest observational astronomers of the nineteenth century expressed serious reservations relative to the idea that the solar surface was illusionary.

Today, much effort continues to be focused on establishing a proper value for the solar radius ([5–12] and references therein). Such reports constitute a clear sign that observational astronomers recognize, at least in practice, the existence of a distinct solar surface. In fact, the measurement of the solar radius not only occupies amateur astronomers, as they map the transits of Mercury and Venus [11, 12], but also attracts the attention of our helioseismologists [5–10]. This is not solely because of the obvious implications for climate change [9]. For theoretical solar physicists, any variation in the dimensions of the Sun would have severe consequences with respect to the gaseous models [5–10]. The latter would be hard-pressed to account for fluctuations in radius. This helps to account for the reassurance experienced when the solar radius is perceived as constant [5–7].

Nonetheless, the solar radius has not definitively been established as fixed. Values obtained in the past thirty years range from $958''.54 \pm 0''.12$ to $960''.62 \pm 0''.02$ (see [10] for a complete table). In 1980, Irwin Shapiro argued that the solar radius had not decreased over time [13]. Currently, these issues cause little debate, though cyclical variations continue to be gently questioned (see [10–13] and references therein).

Perhaps the most interesting aspect of solar radius determinations remains the increased precision of the measurements over the years. Emilio et al. estimate the solar radius at $960''.12 \pm 0''.09$ [10]. This corresponds to 65 km for a radius of more than half a million kilometers (696,342 km) – an error of better than 1 part in 10,000. Others report errors on the order of $0''.02$ [10], a relatively tiny distance of less than 15 km – an error of only 2 parts in 100,000. This precision argues strongly for a distinct solar surface and the existence of a condensed solar body. It is inconceivable that a gaseous Sun would be able to create such a defined "optical illusion". The gaseous solar models argue for smoothly varying density changes, even in the region of the photosphere. As a result, the extreme precision of the solar radius determinations in the visible range, along with previous arguments for a distinct solar surface [3], constitute the twenty-first line of evidence that the Sun is condensed matter.

Additional Note

Chapman et al. [14] have recently reported variability in the Sun's diameter in association with the solar cycle. As previously mentioned, this is a topic of interest to many, though it is only quietly pursued [15]. Variations in the solar diameter with the activity cycle could produce changes in total solar irradiance, beyond the effects produced by sunspots and faculae [16, 17]. While the question of varying solar radius has not been resolved, such phenomena could be accounted for by invoking exfoliative forces within the liquid metallic hydrogen model of the Sun [18]. Exfoliation would be characterized by the production of gases within the condensed solar structure, potential resulting in an expansion of the solar radius. In sharp contrast, changes in radius remain essentially insurmountable within the context of the gaseous models.

Dedication

This work is dedicated to my eldest son, Jacob.

Submitted on: January 25, 2013 / Accepted on: January 25, 2013
First published in online on: January 28, 2013
Additional note on: February 5, 2013

References

1. Keeler J.E. Schmidt's theory of the Sun. *Astrophys. J.*, 1895, v. 1, 178–179.
2. Campbell W.W. James Edward Keeler. *Astrophys. J.*, 1900, v. 12, no. 4, 239–253.
3. Robitaille P.M. On the Presence of a Distinct Solar Surface: A Reply to Hervé Faye. *Progr. Phys.*, 2011, v. 3, 75–78.
4. Hale G.E. Notes on Schmidt's theory of the Sun. *Astrophys. J.*, 1895, v. 2, 69–74.
5. Emilio M., Kuhn J.R., Bush R.I. and Scherrer P. On the constancy of the solar diameter. *Astrophys. J.*, 2000, v. 543, 1007–1010.
6. Kuhn J.R., Bush R.I, Emilio M. and Scherrer P.H. On the constancy of the solar diameter — II. *Astrophys. J.*, 2004, v. 613, 1241–1252.
7. Bush R.I, Emilio M., and Kuhn J.R. On the constancy of the solar diameter — III. *Astrophys. J.*, 2010, v. 716, 1381–1385.
8. Lefebvre S., Bertello L., Ulrich R.K., Boyden J.E., and Rozelot J.P. Solar Radius Measurements at Mount Wilson Observatory. *Astrophys. J.*, 2006, v. 649, 444–451.
9. Raponi A., Sigismondi C., Guhl K, Nugent R. and Tegtmeier A. The measurement of solar diameter and limb darkening function with eclipse observations. *Solar Phys.*, 2012, v. 278, 269–283.
10. Emilio M., Kuhn J.R., Bush R.I, and Scholl I.F. Measuring the solar radius from space during the 2003 and 2006 Mercury transits. *Astrophys. J.*, 2012, v. 750, 135(8 pages).
11. Xie W., Sigismondi C., Wang X. and Tanga P. Venus transit, aureole and solar diameter. In: *"Solar Astrophysical Dynamos and Magnetic Activity"*, *Proc. IAU Symp.*, No. 294, 2013.
12. Sigismondi C. Solar diameter, eclipses and transits: The importance of ground based observations. *Mem. S. A. It.*, 2013, v. 86, 100–103.
13. Shapiro I.I. Is the Sun Shrinking? *Science*, 1980, v. 208, no. 4439, 51–53.
14. Chapman G.A., Dobias J.J. and Walton S.R. On the variability of the apparent solar radius. *Astrophys. J.*, 2008, v. 681, no. 2, 1698–1702.
15. Rozelot J.P. and Damiani C. Rights and wrongs of the temporal solar radius variability. *Eur. Physical J. H*, 2012, v. 37, no. 5, 709–743.
16. de Toma G., White O.R., Chapman G.A., Walton S.R., Preminger D.G., Cookson A.M., Harvey K.L. Differences in the Sun's radiative output in cycles 22 and 23. *Astrophys. J. Letters*, 2001, v. 549, no. 1, L131–L134.
17. Walton S.R., Preminger D.G. and Chapman G.A. The contribution of faculae and network to long-term changes in the total solar irradiance. *Astrophys. J.*, 2003, v. 590, no. 2, 1088–1094.
18. Robitaille J.C. and Robitaille P.M. Liquid Metallic Hydrogen III. Intercalation and lattice exclusion versus gravitational settling and their consequences relative to internal structure, surface activity, and solar winds in the Sun. *Progr. Phys*., 2013, v. 2, 87–97.

LETTERS TO PROGRESS IN PHYSICS

Commentary on the Liquid Metallic Hydrogen Model of the Sun: Insight Relative to Coronal Holes, Sunspots, and Solar Activity

Pierre-Marie Robitaille

Department of Radiology, The Ohio State University, 395 W. 12th Ave, Columbus, Ohio 43210, USA.
robitaille.1@osu.edu

While mankind will always remain unable to sample the interior of the Sun, the presence of sunspots and coronal holes can provide clues as to its subsurface structure. Insight relative to the solar body can also be gained by recognizing that the Sun must exist in the condensed state and support a discrete lattice structure, as required for the production of its continuous spectrum. In this regard, the layered liquid metallic hydrogen lattice advanced as a condensed model of the Sun (Robitaille P.M. Liquid Metallic Hydrogen: A Building Block for the Liquid Sun. *Progr. Phys.*, 2011, v. 3, 60–74; Robitaille P.M. Liquid Metallic Hydrogen II: A Critical Assessment of Current and Primordial Helium Levels in Sun. *Progr. Phys.*, 2013, v. 2, 35–47; Robitaille J.C. and Robitaille P.M. Liquid Metallic Hydrogen III. Intercalation and Lattice Exclusion Versus Gravitational Settling and Their Consequences Relative to Internal Structure, Surface Activity, and Solar Winds in the Sun. *Progr. Phys.*, 2013, v. 2, in press) provides the ability to add structure to the solar interior. This constitutes a significant advantage over the gaseous solar models. In fact, a layered liquid metallic hydrogen lattice and the associated intercalation of non-hydrogen elements can help to account for the position of sunspots and coronal holes. At the same time, this model provides a greater understanding of the mechanisms which drive solar winds and activity.

As the laws of a liquid are different from those of a gas, a liquid star will behave differently from a gaseous star, and before we can predict the behaviour of a star we must know the state of the matter composing it.

James Hopwood Jeans, 1928 [1]

Coronal holes are strange entities, in part due to their sparse nature [2, 3]. At first glance, they seem to offer little of value with respect to our understanding of the Sun. What can be gained from "looking into a hole"? Within the context of the liquid hydrogen model of the Sun (see [4–10] and references therein), there is a great deal to be learned.

In the broadest terms, coronal holes can be described as follows: *"Coronal holes are regions of low-density plasma on the Sun that have magnetic fields that open freely into interplanetary space. During times of low activity, coronal holes cover the north and south polar caps of the Sun. During more active periods, coronal holes can exist at all solar latitudes, but they may only persist for several solar rotations before evolving into a different magnetic configuration. Ionized atoms and electrons flow along the open magnetic fields in coronal holes to form the high speed component of the solar wind"* [2]. When the Sun is quiet, coronal holes appear to be *"anchored"* onto the polar regions of solar surface (see Fig. 1): *"coronal holes, in fact, appear to display rigid rotation as if they are attached to the solar body"* [11, p. 24].

The anchoring of coronal holes to the solar surface can be viewed as the twenty-second line of evidence that the Sun is comprised of condensed matter. The other lines of evidence have already been published (see [4–10] and references therein). Rigid rotation and anchoring cannot be easily explained using the gaseous solar models. As a result, the anchoring of coronal holes is best understood in the context of a condensed solar model.

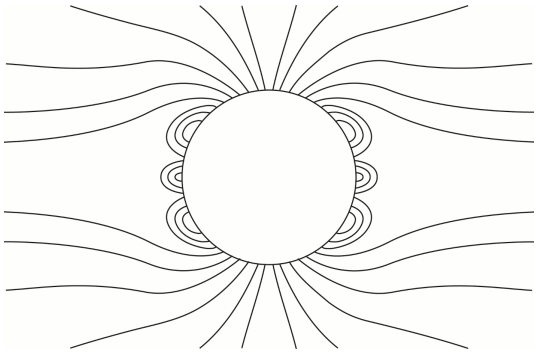

Fig. 1: Schematic representation of coronal holes over the polar caps of a quiet Sun. This figure is an adaptation based on Fig. 2 in [2].

In order to comprehend why the Sun possesses coronal holes, it is best to turn to the lattice configuration of the solar material. Robitaille and Robitaille [7] have recently advanced the hypothesis that the Sun is comprised of liquid metallic hydrogen, wherein protons are arranged in layered hexagonal planes and all other atoms exist in intercalate layers located between the hydrogen planes. Such a structure has been based

on the need to properly explain the thermal emission of the Sun [5], while at the same time, taking into account the structural tendencies of layered materials such as graphite [7].

Within the intercalation compounds of graphite, elemental diffusion orthogonal to the hexagonal carbon planes is hindered, while rapid diffusion can occur in the intercalate regions between the planes (see Fig. 2 in [7]). The same tendencies have been inferred to exist within the liquid metallic hydrogen lattice of the Sun: elemental diffusion is restricted in the direction orthogonal to the hexagonal proton planes and is greatly facilitated within each intercalate layer [7].

Hence, in order to explain the existence of coronal holes, the hexagonal liquid metallic hydrogen lattice of the Sun must be placed in a direction which is orthogonal to the solar surface at the poles. This would explain why the expulsion of ions and electrons from the Sun is facilitated. The subsurface orthogonal placement of the liquid metallic hydrogen hexagonal planes thus accounts for the origin of fast solar winds in these regions. During the quiet periods of the solar cycle, the relative orientation of the hydrogen lattice at the poles forms conduits to drive non-hydrogen elements out of the solar body. As a result, the travel time from the solar core to the surface may well be extremely brief. Given a solar radius of ∼696,342 km (see [10] and references therein) and a fast solar wind of 800 km/s [2], an atom could conceivably leave the solid core of the Sun and escape at the level of the photosphere on the poles in only fifteen minutes.

Nonetheless, during the quiet period of the solar cycle, the equatorial regions of the Sun are unable to sustain fast solar winds. This is likely to occur because the hexagonal layers of liquid metallic hydrogen are parallel to the solar surface in this region. Such an arrangement would restrict the free diffusion of elements from the solar body near the equator, resulting in the absence of fast solar winds. Only the slow component of the solar wind would be observed, precisely because of restricted diffusion of the elements across the hexagonal hydrogen planes [7]. As a result, the concentrations of non-hydrogen elements in the equatorial region of the interior would increase. Eventually, the Sun would become active in order to finally expel these elements from the hydrogen lattice, as was previously stated [7].

Sunspots would be created as hexagonal hydrogen layers are propelled through the solar surface by the force of underlying non-hydrogen elements which have now entered the gaseous phase [7]. This has been illustrated in Fig. 2. Note how the "buckling" of metallic hydrogen could result in the simultaneous formation of two sunspots of opposite polarity (Fig 2, as is usually observed), or of a single sunspot (Fig. 3, as is sometimes observed). Such as scenario also explains why the Sun has relatively "erratic" field lines. These constitute simple extensions of a metallic hydrogen lattice whose internal orientation can be highly variable.

The existence of coronal holes has implications relative to the density of the solar atmosphere. Currently, the gaseous

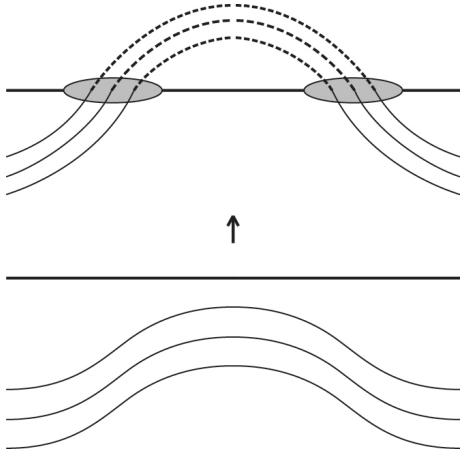

Fig. 2: Schematic representation of the appearance of a pair of sunspots on an active solar surface. The horizontal thick line illustrates the location of the photosphere, the thin lines the layers of metallic hydrogen, and the dashed lines the magnetic field. The two shaded circles outline the position of sunspots. In the lower portion of the figure, the layers of metallic hydrogen are below the level of the photosphere, but are being pushed up by intercalate elements which have entered the gas phase [7]. In the upper portion of the figure, the layers of metallic hydrogen have now broken through the photospheric level. The two sunspots are being linked solely by magnetic field lines, as the metallic hydrogen which once contained them has vaporized into the solar atmosphere. This figure is an adaptation based on Fig. 22 in [12].

solar models are used to assign photospheric and chromospheric densities on the order of 10^{-7} g/cm^3 and 10^{-12} g/cm^3, respectively [12]. In contrast, within the context of the liquid metallic hydrogen model, a photospheric density of ∼1 g/cm^3 is invoked [4–10].

At the same time, the presence of coronal holes directly suggests that chromospheric and coronal densities cannot be spherically uniform. When the Sun is quiet, coronal and chromospheric densities should be lower at the poles and possibly much higher at the equator. Fast solar winds do not typically exist in the equatorial region of the quiet Sun. In fact, it appears that the presence of magnetic field lines restrict the outward movement of ions and electrons away from the solar surface under such conditions. Such realities, when combined with the enormous mass of the Sun, suggest that, contrary to the gaseous solar models, the density of the chromosphere, in the equatorial regions of the quiet Sun, may be many orders of magnitude higher than currently believed. It would be reasonable to suggest that atmospheric densities just above the photospheric layer might far surpass those currently associated with the density of the Earth's atmosphere at sea level. This highlights the problems with extracting densities from regions of the solar atmosphere which are clearly not in local thermal equilibrium, as previously discussed [6].

The liquid metallic hydrogen model [5–7] provides an ex-

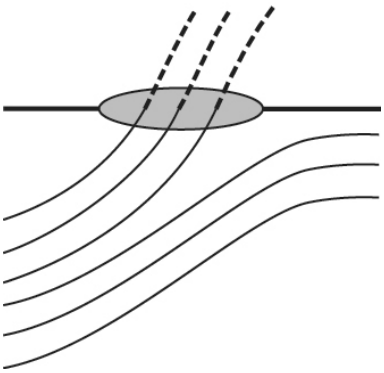

Fig. 3: Schematic representation of a single sunspot on a quiet Sun as in Fig. 2. In this figure, the layers of metallic hydrogen are below the level of the photosphere, but are being pushed up by an adjacent layer of metallic hydrogen which in turn has been displaced by intercalate elements which have entered the gas phase [7]. The sunspot is characterized by strong open magnetic field lines, as the metallic hydrogen which once contained them has vaporized into the solar atmosphere.

cellent framework through which solar activity can be understood. Over the course of the solar cycle, the Sun is alternatively degassing the poles and then the equator. It does so through the orientation of its liquid metallic hydrogen lattice. When the Sun is quiet, much of its interior is constantly being degassed through the action of the fast solar winds exiting at the poles. During this time, degassing is restricted over equatorial regions. Eventually, the Sun becomes active. This change in state is directly associated with degassing the solar interior in the regions of the equator. This helps to explain why sunspots and high energy flares are always restricted to the lower latitudes. They occur in order to degas the equatorial regions of the solar interior. Such a problem does not occur at the poles, since, during the quiet solar period, those internal regions are constantly being degassed by the fast solar winds.

In the end, how the liquid metallic hydrogen layers are oriented within the solar interior reveals a great deal with respect to the formation of sunspots, coronal holes, and measures of solar activity. The magnetic field lines that are observed above the photosphere are a direct consequence of this orientation. Conversely, in the gaseous models of the Sun, the origin of magnetic field lines, coronal holes, sunspots, flares, coronal mass ejections, prominences, and fast or slow solar winds remain areas of profound mystery. This is precisely because these models can offer no structural support for the existence of these phenomena. In order to begin to understand the Sun, structure is required. The continuous solar spectrum requires a lattice for formation. The ideal lattice would resemble the layered one adopted by graphite, as dictated by the needs of thermal emission. Wigner and Huntington have already proposed that metallic hydrogen could adopt a similar lattice [13], creating an ideal structural foundation for the Sun. Furthermore, layered materials, which mimic graphite in their structure, should be prone to forming intercalate regions, as a consequence of lattice exclusion forces [7]. In this regard, the author believes that lattice exclusion, as first postulated by Joseph Christophe Robitaille, along with the formation of intercalate regions within layered metallic hydrogen [7], constitutes the central thesis for understanding solar structure and activity.

Acknowledgment

Luc Robitaille is acknowledged for the preparation of figures.

Dedication

Dedicated to Dmitri Rabounski and Larissa Borissova in fond memory of many scientific discussions on the Sun.

Submitted on: January 25, 2013 / Accepted on: January 25, 2013
First published in online on: January 28, 2013

References

1. Jeans J.H. Astronomy and cosmogony. Cambridge University Press, Cambridge, U.K., 1928, p. 2.
2. Cranmer S.R. Coronal holes. In: *Encyclopedia of Astronomy and Astrophysics*, (Paul Murdin, Ed.), Institute of Physics Publishing, Bristol, 2001, v. 1, p. 496–501.
3. Cranmer S.R. Coronal holes. *Living Rev. Solar Phys.*, 2009, v. 6, 3–66.
4. Robitaille P.M. A high temperature liquid plasma model of the Sun. *Progr. Phys.*, 2007, v. 1, 70–81 (also in arXiv: astro-ph/0410075).
5. Robitaille P.M. Liquid Metallic Hydrogen: A Building Block for the Liquid Sun. *Progr. Phys*., 2011, v. 3, 60–74.
6. Robitaille P.M. Liquid Metallic Hydrogen II: A Critical Assessment of Current and Primordial Helium Levels in Sun. *Progr. Phys.*, 2013, v. 2, 35–47.
7. Robitaille J.C. and Robitaille P.M. Liquid Metallic Hydrogen III. Intercalation and Lattice Exclusion Versus Gravitational Settling and Their Consequences Relative to Internal Structure, Surface Activity, and Solar Winds in the Sun. *Progr. Phys*., 2013, v. 2, in press.
8. Robitaille P.M. Commentary relative to the distribution of gamma-ray flares on the Sun: Further evidence for a distinct solar surface. *Progr. Phys.*, 2013, v. 2, L1-L2.
9. Robitaille P.M. Commentary Relative to the Seismic Structure of the Sun:
Internal Rotation, Oblateness, and Solar Shape. *Progr. Phys.*, 2013, v. 2, L3-L4.
10. Robitaille P.M. Commentary on the Radius of the Sun:
Optical Illusion or Manifestation of a Real Surface? *Progr. Phys.*, 2013, v. 2, L5-L6.
11. Chitre S.M. Overview of solar physics. In: *Lectures on Solar Physics* (H.M. Antia, A. Bhatnagar and R. Ulmschneider, Eds.), Springer, Berlin, 2003, p. 1–26.
12. Bhatnagar A. Instrumentation and observational techniques in solar astronomy. In: *Lectures on Solar Physics* (H.M. Antia, A. Bhatnagar and R. Ulmschneider, Eds.), Springer, Berlin, 2003, p. 27–79.
13. Wigner E. and Huntington H.B. On the possibility of a metallic modification of hydrogen. *J. Chem. Phys.*, 1935, v. 3, 764–70.

LETTERS TO PROGRESS IN PHYSICS

Commentary on the Liquid Metallic Hydrogen Model of the Sun II. Insight Relative to Coronal Rain and Splashdown Events

Pierre-Marie Robitaille

Department of Radiology, The Ohio State University, 395 W. 12th Ave, Columbus, Ohio 43210, USA.
robitaille.1@osu.edu

Coronal rain represents blobs of solar material with a width of ∼300 km and a length of ∼700 km which are falling from the active region of the corona towards the solar surface along loop-like paths. Conversely, coronal showers are comprised of much larger bulks of matter, or clumps of solar rain. Beyond coronal rain and showers, the expulsion of solar matter from the surface, whether through flares, prominences, or coronal mass ejections, can result in massive disruptions which have been observed to rise far into the corona, return towards the Sun, and splashdown onto the photosphere. The existence of coronal rain and the splashdown of mass ejections onto the solar surface constitute the twenty-third and twenty-fourth lines of evidence that the Sun is condensed matter.

As the laws of a liquid are different from those of a gas, a liquid star will behave differently from a gaseous star, and before we can predict the behaviour of a star we must know the state of the matter composing it.

James Hopwood Jeans, 1928 [1]

The presence of coronal rain within the active atmosphere of the Sun has been recognized for less than a decade [2–5]. Coronal rain corresponds to *"cool and dense matter and not waves"* [5]. It appears to be *"ubiquitous"* and *"composed of a myriad of small blobs, with sizes that are, on average 300 km in width and 700 km in length"* [5]. When it aggregates, coronal rain can lead to larger clumps called *"showers"* [5]. Their rate of descent towards the solar surface can approach 120 km s^{-1}. However, such rates of descent are inferior to those inferred from the Sun's gravitational field, suggesting that they are restricted in their downward motion by gas pressure in the underlying solar atmosphere [5]. These findings are incongruent with the idea that the density of the chromosphere is in the 10^{-12} g/cm^3 range, as currently advanced by the gaseous solar models [6]. Such densities would be associated with very good vacuums on Earth. As such, it does not seem reasonable, based on these findings, that the chromospheric densities associated with the gaseous models can be correct [7]. At the same time, theoretical models relative to coronal rain now rely on *"heating and condensation cycles"* [4, 5], despite the fact that the gaseous models of the Sun preclude all material condensation. In the end, it remains more plausible to account for the behavior of coronal rain by invoking true condensation, as seen in the liquid metallic model of the Sun [7]. This constitutes the twenty-third line of evidence that the Sun is comprised of condensed matter (see [7] and references therein for the others).

In addition to coronal rain, the mass ejection event, witnessed on June 7, 2011, was particularly instructive relative to the nature of the Sun [8, 9]. On that day, a tremendous disruption occurred on the solar surface which projected material well into the corona, prior to its subsequent descent back onto the Sun. Upon striking the solar body, the multiple points of impact immediately brightened – revealing clear and distinct surface behavior on the photosphere [9]. Such visualizations highlight that the solar surface is not an optical illusion, but, indeed, acts as a real surface. Such "splashdowns" constitute the twenty-fourth line of evidence that the solar body is comprised of condensed matter. In addition, they provide complementary evidence that flares, prominences, and coronal mass ejections are also characterized by the existence, at least in part, of condensed matter.

Impressive disruptions of the solar surface have also been associated with comets, although intial analysis apparently revealed that such events were not associated with the impact of such objects onto the photosphere [10]. In the end, additional study may well reveal that comets have the ability to disrupt the solar surface, either directly through impact or indirectly by disrupting magnetic field lines above the surface.

Such visualizations highlight that the solar surface is not an optical illusion. It appears and behaves as a true liquid surface. In addition, coronal rains and mass ejection splashdowns indicate that the outer atmosphere of the Sun can support localized regions of condensed matter.

Acknowledgment

The May 10–11, 2011 solar event [10] was first brought to the author's attention by Patrice Robitaille.

Dedication

This work is dedicated to my brother, Patrice, in profound gratitude for many years of support and encouragement.

Submitted on: January 31, 2013 / Accepted on: January 31, 2013
First published in online on: February 2, 2013

References

1. Jeans J.H. Astronomy and cosmogony. Cambridge University Press, Cambridge, U.K., 1928, p. 2.
2. De Groof A., Berghmans D., van Driel-Gesztelyi L. and Poedts S. Intensity variations in EIT shutterless mode: Waves or Flows? *Astron. Astrophys.*, 2004, v. 415, no. 3, 1141–1151.
3. De Groof A., Bastiaensen C., Müller D.A.N., Berghmans D., and Poedts S. Detailed comparison of downflows seen both in EIT 30.4 nm and Big Bear Hα movies. *Astron. Astrophys.*, 2005, v. 443, no. 1, 319–328.
4. Müller D.A.N., De Groof A., Hansteen V.H. and Peter H. High-speed coronal rain. *Astron. Astrophys.*, 2005, v. 436, no. 3, 1067–1074.
5. Antolin R., Vissers G. and van der Voort L.R. On-Disk coronal rain. *Solar Phys.*, 2012, v. 280, no. 2, 457–474.
6. Bhatnagar A. Instrumentation and observational techniques in solar astronomy. In: *Lectures on Solar Physics* (H.M. Antia, A. Bhatnagar and R. Ulmschneider, Eds.), Springer, Berlin, 2003, p. 27–79.
7. Robitaille P.M. Commentary on the liquid metallic hydrogen model of the Sun: Insight relative to coronal holes, sunspots, and solar activity. *Progr. Phys.*, 2013, v. 2, L7–L9.
8. Moskowitz C. Gargantuan Sun explosion rocks astronomers. *SPACE.com* (created June 9, 2011 at 12:34 PM ET – accessed online on January 29, 2011).
9. NASA/SDO/Helioviewer.org [2011/06/07 04:00:00 to 11:00:00 UTC]. Observed well using 5 min frames SDO AIA 304. These events have been captured in video format and posted numerous times online: e.g.Newsflash *Skywatch Media* youtube.com/watch?v=aQICN0BV1Aw; Phil Plait *Bad Astronomy Blog* youtube.com/watch?v=Hyi4hjG6kDM; youtube.com/watch?v=Cjd6TQiTRAM; youtube.com/watch?v=6h1EJtmUTeM. (Accessed online on January 29, 2013).
10. SOHO NASA/ESA [2011/05/10 20:00:00 to 2011/05/11 08:00:00 UTC]. These events have been captured in video format and displayed online: e.g. youtube.com/watch?v=igeBrSGk5FA; *Russia Today* youtube.com/watch?NR=1&v=Mat4dWpszoQ&feature=fvwp. (Accessed online on January 29, 2013).

LETTERS TO PROGRESS IN PHYSICS

Commentary on the Liquid Metallic Hydrogen Model of the Sun III. Insight into Solar Lithium Abundances

Pierre-Marie Robitaille

Department of Radiology, The Ohio State University, 395 W. 12th Ave, Columbus, Ohio 43210, USA.
robitaille.1@osu.edu

The apparent depletion of lithium represents one of the greatest challenges to modern gaseous solar models. As a result, lithium has been hypothesized to undergo nuclear burning deep within the Sun. Conversely, extremely low lithium abundances can be easily accounted for within the liquid metallic hydrogen model, as lithium has been hypothesized to greatly stabilize the formation of metallic hydrogen (E. Zurek et al. A little bit of lithium does a lot for hydrogen. *Proc. Nat. Acad. Sci. USA*, 2009, v. 106, no. 42, 17640–17643). Hence, the abundances of lithium on the solar surface can be explained, not by requiring the nuclear burning of this element, but rather, by suggesting that the Sun is retaining lithium within the solar body in order to help stabilize its liquid metallic hydrogen lattice. Unlike lithium, many of the other elements synthesized within the Sun should experience powerful lattice exclusionary forces as they are driven out of the intercalate regions between the layered liquid metallic hydrogen hexagonal planes (Robitaille J.C. and Robitaille P.M. Liquid Metallic Hydrogen III. Intercalation and Lattice Exclusion Versus Gravitational Settling and Their Consequences Relative to Internal Structure, Surface Activity, and Solar Winds in the Sun. *Progr. Phys*., 2013, v. 2, in press). As for lithium, its stabilizing role within the solar interior helps to account for the lack of this element on the surface of the Sun.

As the laws of a liquid are different from those of a gas, a liquid star will behave differently from a gaseous star, and before we can predict the behaviour of a star we must know the state of the matter composing it.

James Hopwood Jeans, 1928 [1]

Solar lithium abundance [2], as determined at the photospheric level, are reduced ~140 fold when compared to meteorites [3]. Such a paucity of lithium has presented a challenge for the gaseous models of the stars, as they attempt to account for the relative absence of this element on the solar surface [2,3]. Consequently, solar scientists hypothesized that lithium is being burned deep within the convection zone [2,3]. Lithium is thought to be easily destroyed [^{7}Li(p,α)^{4}He] at temperatures above 2.6 x 10^6 K [4]. Mild mixing of lithium also helps to account for the surface depletion [4–6]. In this regard, it has been postulated that *"stars that host planets experience more mixing in their internal environment"* [7]. As a result, those who adhere to the gaseous models have proposed that greater lithium depletion occurs in stars that have orbiting planets [8], although such claims have been refuted [9]. Nonetheless, such works [7, 9] highlight the significance of the solar lithium abundance problem in astrophysics. In this regard, solar lithium abundances might be better understood within the context of the liquid metallic hydrogen model of the Sun [10–13].

Along with Neil Ashcroft, Eva Zurek and her coworkers recently advanced [14] that lithium could greatly stabilize the formation of metallic hydrogen [15, 16]. This finding has tremendous implication relative to understanding the fate of lithium within the Sun, if indeed, the solar matrix is comprised of liquid metallic hydrogen [10–13].

When the Sun was hypothesized to be built from liquid metallic hydrogen, it was important that the resulting lattice adopt a layered structure similar to graphite in order to properly account for thermal emission [11]. Thus, it was fortunate that Wigner and Huntington [15] had said that metallic hydrogen could exist in a layered lattice resembling graphite. At the same time, since graphite was known to form intercalation compounds, the extension of such chemistry to the layered form of metallic hydrogen proved natural [13]. Therefore, it was thought that the Sun would maintain the integrity of its layered hexagonal hydrogen lattice and associated conduction bands, by permitting non-hydrogen elements to reside within intercalation zones [13]. In addition, since the intercalation compounds of graphite are known to undergo exfoliative processes wherein intercalate atoms are driven out of the graphitic structure, the same mechanism was applied to the Sun [13]. Solar activity became linked to lattice exclusion and the associated expulsion of non-hydrogen atoms from the solar interior [13]. Nonetheless, it was already recognized [11] that lithium should stabilize the metallic hydrogen lattice. As a result, unlike the case for most elements, the Sun should not be working to expel lithium. Such a scenario elegantly accounts for the significant reductions in lithium abundances observed on the surface of the Sun while, at the same time,

permitting elevated lithium levels in meteorites, or other objects, which have been first synthesized within the stars. Conversely, the idea that lithium is being burned preferentially within the stars, as proposed by the gaseous models, makes it difficult to account for elevated lithium levels elsewhere in the astrophysical world. Herein lies the merit of sequestering lithium within the solar body and permitting it to participate in nuclear reactions, without preferential burning, in the context of the liquid metallic hydrogen model [10–13].

Dedication

This work is dedicated to my youngest son, Luc.

Submitted on: January 31, 2013 / Accepted on: January 31, 2013
First published in online on: February 2, 2013

References

1. Jeans J.H. Astronomy and cosmogony. Cambridge University Press, Cambridge, U.K., 1928, p. 2.
2. Grevesse N. Solar abundances of lithium, beryllium, and boron. *Solar Phys.*, 1968, v. 5, 159–180.
3. Wiens R.C., Bochsler P., Burnett D.S. and Wimmer-Schweingruber R.F. Solar and solar-wind isotopic compositions. *Earth Plan. Sci. Let.*, 2004, v. 222, 697–712.
4. Uitenbroek H. The effect of photospheric granulation on the determination of the lithium abundance in solar-type stars. *Astrophys. J.*, 1998, v. 498, 427–440.
5. Schatzman E. Turbulent transport and lithium destruction in main sequence stars. *Astron. Astrophys.*, 1977, v. 56, 211–218.
6. Pinsonneault M.H. Rotational mixing and lithium depletion. In: *Light Elements in the Universe: Proc. IAU Symposium*, 2009, v. 268, 375-380.
7. Pinsonneault M.H. A fossil record for exoplanets. *Nature*, 2009, v. 462, no. 7270, 168–169.
8. Israelian G., Delgado Mena E., Santos N.C., Sousa S.G., Mayor M., Udry S., Dominguez Cerdena C., Rebolo R. and Randich S. Enhanced lithium depletion in Sun-like stars with orbiting planets. *Nature*, 2009, v. 462, no. 7270, 189–191.
9. Baumann P., Ramirez I., Melendez J., Aslund M. and Lind K. Lithium depletion in solar-like stars: no planet connection. *Astron. Astrophys.*, 2010, v. 519, A87(11 pages).
10. Robitaille P.M. A high temperature liquid plasma model of the Sun. *Progr. Phys.*, 2007, v. 1, 70–81 (also in arXiv: astro-ph/0410075).
11. Robitaille P.M. Liquid Metallic Hydrogen: A Building Block for the Liquid Sun. *Progr. Phys.*, 2011, v. 3, 60–74.
12. Robitaille P.M. Liquid Metallic Hydrogen II: A Critical Assessment of Current and Primordial Helium Levels in Sun. *Progr. Phys.*, 2013, v. 2, 35–47.
13. Robitaille J.C. and Robitaille P.M. Liquid Metallic Hydrogen III. Intercalation and Lattice Exclusion Versus Gravitational Settling and Their Consequences Relative to Internal Structure, Surface Activity, and Solar Winds in the Sun. *Progr. Phys.*, 2013, v. 2, in press.
14. Zurek E., Hoffmann R., Ashcroft N.W., Oganov A.R., Lyakhov A.O. A little bit of lithium does a lot for hydrogen. *Proc. Nat. Acad. Sci. USA* USA, 2009, v. 106, no. 42, 17640—17643.
15. Wigner E. and Huntington H.B. On the possibility of a metallic modification of hydrogen. *J. Chem. Phys.*, 1935, v. 3, 764–70.
16. McMahon J.M., Morales M.A., Pierleoni C. and Ceperley D.M. The properties of hydrogen and helium under extreme conditions. *Rev. Mod. Phys.*, 2012, v. 84, 1607–1653.

The Journal on Advanced Studies in Theoretical and Experimental Physics, including Related Themes from Mathematics

PROGRESS IN PHYSICS

A quarterly issue scientific journal, registered with the Library of Congress (DC, USA). This journal is peer reviewed and included in the abstracting and indexing coverage of: Mathematical Reviews and MathSciNet (AMS, USA), DOAJ of Lund University (Sweden), Zentralblatt MATH (Germany), Scientific Commons of the University of St. Gallen (Switzerland), Open-J-Gate (India), Referativnyi Zhurnal VINITI (Russia), etc.

Electronic version of this journal:
http://www.ptep-online.com

Postal Address

Department of Mathematics and Science,
University of New Mexico,
705 Gurley Ave., Gallup, NM 87301, USA

ISSN: 1555-5534 (print)
ISSN: 1555-5615 (online)

Standard Address Number: 297-5092
Printed in the United States of America

JULY 2013 **VOLUME 3**

CONTENTS

Continued on the next page ⟶

Information for Authors and Subscribers

Progress in Physics has been created for publications on advanced studies in theoretical and experimental physics, including related themes from mathematics and astronomy. All submitted papers should be professional, in good English, containing a brief review of a problem and obtained results.

All submissions should be designed in LaTeX format using *Progress in Physics* template. This template can be downloaded from *Progress in Physics* home page. Abstract and the necessary information about author(s) should be included into the papers. To submit a paper, mail the file(s) to the Editor-in-Chief.

All submitted papers should be as brief as possible. Letters related to the publications in the journal or to the events among the science community can be applied to the section *Letters to Progress in Physics*.

All that has been accepted for the online issue of *Progress in Physics* is printed in the paper version of the journal. To order printed issues, contact the Editors.

This journal is non-commercial, academic edition. It is printed from private donations. (Look for the current author fee in the online version of the journal.)

CONTENTS

(continued from the previous page)

LETTERS:

On the Propagation of Light in an Expanding Universe

Yuri Heymann

3 rue Chandieu, 1202 Geneva, Switzerland. E-mail: y.heymann@yahoo.com

The equation of the propagation of light in an expanding Universe is derived based on the definition of comoving distances. A numerical method is proposed to solve this equation jointly with the Friedmann equation. As the equation of the propagation of light in an expanding Universe defines a horizon of the visible Universe, this puts a constraint on cosmological models in order to be consistent with an upper limit for redshifts observed from galaxies. This puzzle is challenging current expansionist cosmological models.

1 Introduction

Euclidean Distances were introduced in [1] in order to derive the galactic density profile which is the evolution of galactic density over time. We define the Euclidean Distance as the equivalent distance that would be traversed by a photon between the time it is emitted and the time it reaches the observer if there were no expansion of the Universe. The comoving distance is the distance between two points measured along a path defined at the present cosmological time. The comoving distance between objects moving with the Hubble flow is deemed to remain constant in time. The Euclidean Distance is also the proper distance at the time of emission for a source of light, which is the comoving distance multiplied by the scale factor at the time of emission. From this relationship, the equation of the propagation of light in an expanding Universe is derived.

2 Equation of the propagation of light in an expanding Universe

As the Euclidean Distance is the proper distance at the time light was emitted from a source of light, it is equal to the comoving distance times the scale factor at the time of emission. By convention the scale factor is equal to one at the present time. Therefore, we have

$$y = a(t)\chi\,, \tag{1}$$

and

$$\chi = c\int_{t=T_b-T}^{T_b}\frac{dt}{a(t)}\,, \tag{2}$$

where χ is the comoving distance, y the Euclidean Distance, a the scale factor, T_b the time from the hypothetical big bang (which is the present time), and T the light travel time between observer and the source of light.

By differentiating (1) with respect to time we get:

$$\frac{dy}{dt} = \dot{a}\chi + a\dot{\chi}\,. \tag{3}$$

As $I = \int_{t_1}^{t_2} f(t)\,dt$ leads to $\frac{dI}{dt} = \frac{dt_2}{dt}f(t_2) - \frac{dt_1}{dt}f(t_1)$, from (2) we get:

$$\dot{\chi} = -\frac{c}{a(t)}\,. \tag{4}$$

As $H(t) = \dot{a}/a$, (1) leads to:

$$\dot{a}\chi = y\,H(t)\,, \tag{5}$$

Combining (3), (4) and (5) we get:

$$\frac{dy}{dt} = -c + H(t)\,y\,, \tag{6}$$

where y is the Euclidean Distance between the observer and a photon moving towards the observer.

We have just derived the equation of the propagation of light in an expanding Universe from the definition of comoving distances. This equation defines a horizon of the visible Universe at $\frac{dy}{dt} = 0$.

3 Numerical method to compute Euclidean Distances from the Friedmann equation

Equation (6) can be solved numerically using a discretization method. Let us set $t = T_b - T$ with T_b the hypothetical time since the big bang, and T the light travel time between observer and the photon. Therefore, $dt = -dT$, and (6) can be rewritten as follows:

$$\frac{dy}{dT} = c - H(T)\,y\,. \tag{7}$$

By discretization over small intervals ΔT, (7) leads to:

$$\frac{y_{n+1} - y_n}{\Delta T} = c - H(T_n)\,y_n\,. \tag{8}$$

Therefore, we obtain:

$$y_{n+1} = c\,\Delta T + y_n\,(1 - H(T_n)\,\Delta T)\,, \tag{9}$$

with initial conditions: $y_0 = 0$ and $T_0 = 0$, and $T_{n+1} = T_n + \Delta T$.

The Friedmann equation expresses H as a function of redshift z. We still need a description of H as a function of T in order to solve (9). For this purpose we compute a curve for the light travel time T versus redshift z using the Friedmann equation, with (11). Then we fit an empirical equation for $H(T)$ over the curve $H(z)$ versus T.

The light travel time versus redshift is computed as follows (derived from $dt = da/\dot{a}$):

$$T = c \int_{1/(1+z)}^{1} \frac{da}{\dot{a}} . \quad (10)$$

Because $H = \dot{a}/a$, (10) can be rewritten as follows:

$$T = c \int_{1/(1+z)}^{1} \frac{da}{H\, a} . \quad (11)$$

This integral is solved numerically using a solver such as Matlab.

The Friedmann equation that is used in this problem is as follows:

$$H = H_0 \sqrt{\Omega_R a^{-4} + \Omega_M a^{-3} + \Omega_k a^{-2} + \Omega_\Lambda} , \quad (12)$$

with Ω_R the radiation energy density today, Ω_M the matter density today, Ω_k the spatial curvature density today, and Ω_Λ a cosmological constant for the vacuum energy density today. We may alternatively express H as a function of redshift from cosmological redshift relationship by setting $a = \frac{1}{1+z}$, where the scale factor is equal to unity as the present time.

4 Results and discussion

First let us solve the above problem with the assumptions used in the lambda-cdm model [2]. The radiation energy density is generally considered neglegible, hence $\Omega_R = 0$. The common assumption in the lambda-cdm is that Ω_k is equal to zero, and $\Omega_\Lambda = 1 - \Omega_M$. To obtain a description of H as a function of T, we fit a polynomial function of order six to the $H(z)$ curve, which gives the following empirical formula for $\Omega_M = 0.3$ and $H_0 = 71\, km\, s^{-1}\, Mpc^{-1}$: $H(T) = 0.074663 - 0.049672\, T + 0.056296\, T^2 - 0.021203\, T^3 + 0.0036443\, T^4 - 0.00029054\, T^5 + 0.0000088134\, T^6$, with T in $Glyr$ and $H(T)$ in $Glyr^{-1}$. From the discretization method (9) we obtain an horizon of the visible Universe at redshift $z = 1.6$. A variant of the lambda-cdm model would be to remove the cosmological constant for the vacuum energy density ($\Omega_\Lambda = 0$), and replace this term by the spatial curvature density $\Omega_k = 1 - \Omega_M$. This variant gives almost the same result with a horizon of the visible Universe at redshift $z = 1.5$. On the other hand if H is constant over time, the horizon of the visible Universe would have a redshift that tends to infinity.

The results obtained with the equation we derived for the propagation of light solved jointly with the Friedmann equation are inconsistent with observations as it is common to observe galaxies with redshifts up to 6, and more recently beyond 8.5 [3]. This problem has been raised in the past – the recession velocity of all galaxies with $z \geq 1.5$ has been found to exceed the speed of light in all viable cosmological models [4]. A calculation based on null geodesics using gravitational radius is proposed in [5]. Their hypothesis is that the comoving distance and proper distance do not track the propagation of light through the Hubble flow. The puzzle of the propagation of light in an expanding Universe and the horizon of the visible Universe appears to be an interesting challenge for current expansionist cosmological models.

Submitted on February 16, 2013 / Accepted on February 23, 2013

References

1. Heymann Y. Building galactic density profiles. *Progress in Physics*, 2011, v. 4, 63–67.
2. Wright E. L. A Cosmology Calculator for the World Wide Web. *The Publications of the Astronomical Society of the Pacific*, 2006, v. 118, 1711–1715.
3. Ellis R. S., McLure R. J., Dunlop J. S., Robertson B. E., Ono Y., Schenker M. A., Koekemoer A., Bowler R. A. A., Ouchi M., Rogers A. B., Curtis-Lake E., Schneider E., Charlot S., Stark D. P., Furlanetto S. R., and Cirasuolo M. The abundance of star-forming galaxies in the redshift range 8.5-12: new results from the 2012 Hubble Ultra deep field campaign. *The Astrophysical Journal Letters*, 2013, v. 763, 1–6.
4. Davis T. and Lineweaver C. H. Expanding Confusion: Common Misconceptions of Cosmological Horizons and the Superluminal Expansion of the Universe. *Publications of the Astronomical Society of Australia*, 2004, v. 21, 97–109.
5. Bikwa O., Melia F., and Shevchuk A. Photon Geodesics in FRW Cosmologies. *Monthly Notices of the Royal Astronomical Society*, 2012, v. 421, 3356–3361.

On the Luminosity Distance and the Hubble Constant

Yuri Heymann

3 rue Chandieu, 1202 Geneva, Switzerland. E-mail: y.heymann@yahoo.com

By differentiating luminosity distance with respect to time using its standard formula we find that the peculiar velocity is a time varying velocity of light. Therefore, a new definition of the luminosity distance is provided such that the peculiar velocity is equal to c. Using this definition a Hubble constant $H_0 = 67.3\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ is obtained from supernovae data.

1 Introduction

The luminosity distance is an important concept in cosmology as this is the distance measure obtained from supernovae data using the distance modulus. The standard formula of the luminosity distance is $d_L = (1+z)\,d_M = d_M/a$, where d_L is the luminosity distance and d_M the comoving transverse distance [1, p. 421]. As shown below this definition implies that the peculiar velocity is a time varying velocity of light, and therefore a new definition is proposed where the speed of light is constant.

2 Definition of the luminosity distance and the peculiar velocity from light propagation

From there we will use the notation r_L for the luminosity distance as it represents the radius of a sphere for light propagating from the center which is the point of emission of the light source. The standard formula of the luminosity distance for a flat Universe is as follows:

$$r_L = \frac{\chi}{a}\,, \tag{1}$$

and

$$\chi = c\int_0^{t_0}\frac{dt}{a}\,, \tag{2}$$

where r_L is the luminosity distance, χ the comoving distance, a the scale factor at the time of emission, t the time which is equal to zero at the origin set at the center of the sphere from which light is emitted, and t_0 the time when light reaches the earth.

Let us apply the change of coordinates $T = t_0 - t$, where T is the light travel time between the observer and the photon. Hence, $dt = -dT$, and (2) can be rewritten as follows:

$$\chi = -c\int_T^0\frac{dT}{a} = c\int_0^T\frac{dT}{a}\,. \tag{3}$$

By differentiating (1) with respect to T we get:

$$\frac{dr_L}{dT} = \frac{\dot\chi}{a} - \frac{\dot a}{a^2}\chi\,. \tag{4}$$

As $I = \int_{t_1}^{t_2} f(t)\,dt$ leads to $\frac{dI}{dt} = \frac{dt_2}{dt}\,f(t_2) - \frac{dt_1}{dt}\,f(t_1)$, from (3) we get:

$$\dot\chi = \frac{c}{a}\,. \tag{5}$$

Using (1) we get:

$$\frac{\dot a}{a^2}\chi = \frac{\dot a}{a}\,r_L\,. \tag{6}$$

Because $H = \frac{1}{a}\frac{da}{dt} = -\frac{1}{a}\frac{da}{dT}$, equation (6) can be rewritten as follows:

$$\frac{\dot a}{a^2}\chi = -H\,r_L\,. \tag{7}$$

Combining (4), (5) and (7) we get:

$$\frac{dr_L}{dT} = \frac{c}{a^2} + H\,r_L\,. \tag{8}$$

The term $H\,r_L$ represents the expansion for the radius of our sphere, and $\frac{c}{a^2}$ is the peculiar velocity. From light propagation we see that the standard formula of luminosity distance implies a time varying velocity of light.

A new equation is proposed for the luminosity distance where the peculiar velocity is always equal to c. Considering a sphere of radius r'_L for the propagation of light emitted from a point at the center, and that the sphere inflates over time due to the expansion of the Universe and the velocity of light, we obtain:

$$\frac{dr'_L}{dT} = c + H\,r'_L\,, \tag{9}$$

with boundary condition $r'_L = 0$ at $T = 0$. Where r'_L is the luminosity distance, T the light travel time between emission and reception of the light source, and H the Hubble constant at time T.

3 Solving the equation of the luminosity distance

In this section we assume that the Hubble constant does not vary over time and is always equal to H_o.

By integrating (9) we get:

$$r'_L = \frac{c}{H_0}\,(\exp(H_0\,T) - 1)\,. \tag{10}$$

This equation can be rewritten as follows:

$$T = \frac{1}{H_0}\,\ln\!\left(1 + \frac{H_0}{c}\,r'_L\right). \tag{11}$$

The expression of the light travel time versus redshift is as follows:

$$T = \int_{1/(1+z)}^{1}\frac{da}{H\,a} = \frac{1}{H_0}\,\ln(1+z)\,. \tag{12}$$

By combining (11) and (12) we get:

$$r'_L = \frac{c}{H_0} z \,. \tag{13}$$

4 Calculation of the Hubble constant from supernovae data

Let us compute the Hubble constant from supernovae using the relationship in (13) . In order to compute the luminosity distance we use the redshift adjusted distance modulus provided in [2] which is as follows:

$$m - M = -5 + 5 \log r'_L + 2.5 \log(1 + z) \,. \tag{14}$$

The distance modulus $\mu = m - M$ is the difference between the apparent magnitude m and the absolute magnitude M.

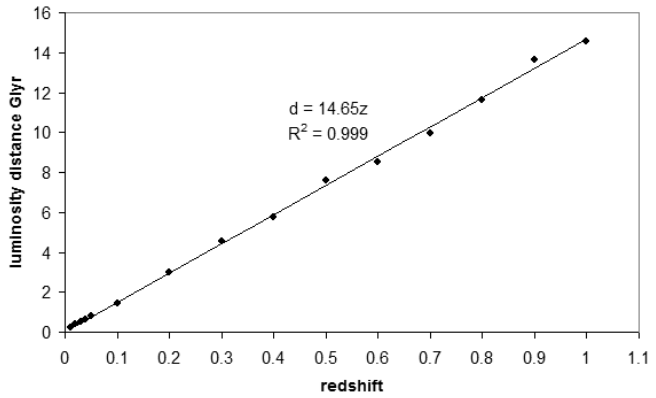

Fig. 1: Luminosity distance in Glyr versus redshift plot for supernovae. Data source: http://supernova.lbl.gov/Union/

In Fig. 1 we have a plot of the luminosity distance versus redshift that was obtained with (14) using supernovae data. This plot is rectilinear with a slope of 14.65 where the luminosity distance is expressed in *Glyr* (billion light years). The Hubble constant which is the inverse of the slope from (13) is equal to $H_0 = 67.3\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$.

5 Conclusion

In this study it has been shown that the standard formula of the luminosity distance implies that the peculiar velocity is a time varying velocity of light. Given our choice for the luminosity distance equation which is based on a peculiar velocity always equal to c, we find that the solution to this equation requires a Hubble constant that does not change over time in order to fit the supernovae data.

Submitted on February 16, 2013 / Accepted on February 23, 2013

References

1. Weinberg S. Gravitation and Cosmology: Principles and applications of the general theory of relativity. John Wiley and Sons, 1972.
2. Heymann Y. Redshift Adjustment to the Distance Modulus. *Progress in Physics*, 2012, v. 1, 6–7.

Electric Dipole Antenna: A Source of Gravitational Radiation

Chifu E. Ndikilar* and Lawan S. Taura†

*Physics Department, Federal University Dutse, Nigeria
†Physics Department, Bayero University Kano, Nigeria
E-mail: ebenechifu@yahoo.com

In this article, the gravitational scalar potential due to an oscillating electric dipole antenna placed in empty space is derived. The gravitational potential obtained propagates as a wave. The gravitational waves have phase velocity equal to the speed of light in vacuum (c) at the equatorial plane of the electric dipole antenna, unlike electromagnetic waves from the dipole antenna that cancel out at the equatorial plane due to charge symmetry.

1 Introduction

Gravitational waves were predicted to exist by Albert Einstein in 1916 on the basis of the General Theory of Relativity. They are usually produced in an interaction between two or more compact masses. Such interactions include the binary orbit of two black holes, a merge of two galaxies, or two neutron stars orbiting each other. As the black holes, stars, or galaxies orbit each other, they send out waves of "gravitational radiation" that reach the Earth. A lot of efforts have been made over the years to detect these very weak waves. In this article, we show theoretically, how the gravitational potential of an electric dipole antenna placed in empty space propagates as gravitational waves.

2 Gravitational radiation from an electric dipole antenna

Recall that an electric dipole antenna is a pair of conducting bodies (usually spheres or rectangular plates) of finite capacitance connected by a thin wire of negligible capacitance through an oscillator. The charges reside on the conducting bodies (electrodes) but may travel from one to the other through the wire. The oscillator causes the charges to be built up on the electrodes such that at any time they are equal and opposite and the variation is sinusoidal with angular frequency ω [1].

Let the electric dipole antenna be represented by a pair of spheres seperated by a distance s with a sinusoidal charge Q as shown in figure 1.

If the total mass of each sphere at any time is M_0 and its radius R_s and assuming an instantaneous mass distribution which varies with the motion of electrons, then at each time t, the mass density ρ_0 is given by

$$\rho_0 = \Lambda_0 + \rho_e \sin \omega t \tag{1}$$

where

$$\Lambda_0 = \frac{M_0}{4\pi R_s^3}$$

and

$$\rho_e = \frac{Nm_e}{4\pi R_s^3}$$

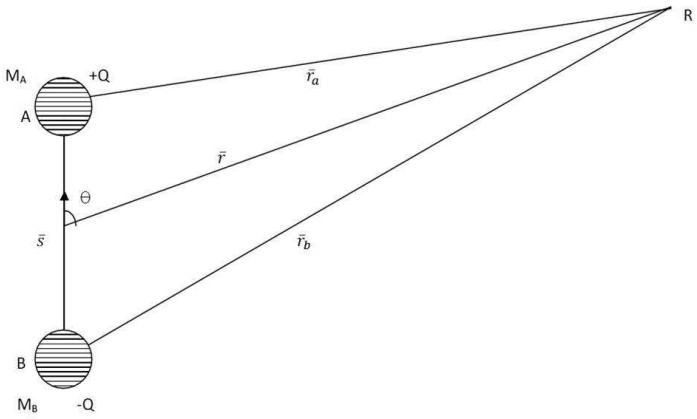

Fig. 1: Amplified diagram of an electric dipole antenna.

where N is the number of electrons moving in the dipole antenna and m_e is the electronic mass. For this mass distribution, the gravitational field equation can be written as [2]

$$\nabla^2\Phi = \begin{cases} 0 & \text{if } r > R \\ 4\pi G\rho_0 & \text{if } r < R \end{cases} \tag{2}$$

Now, consider a unit mass placed at a point R in empty space, far off from the electric dipole as in figure 1, then by Newton's dynamical theory, the gravitational scalar potential Φ at R at any time t can be defined as

$$\Phi(\bar{r}, t) = \frac{GM_a(\bar{r_a}, t)}{|\bar{r} - \bar{r_a}|} + \frac{GM_b(\bar{r_b}, t)}{|\bar{r} - \bar{r_b}|}. \tag{3}$$

To maintain equal and opposite charges at the electrodes, the sinusoidal movement of electrons must be in such a way that the masses of the two spheres are the same and determined at point R to be given by

$$M_a(\bar{r_a}, t) = M_b(\bar{r_b}, t) = M_0 e^{i\omega(t')}. \tag{4}$$

Thus, the gravitational potential at R becomes

$$\Phi(\bar{r}, t) = \frac{GM_0 e^{i\omega(t')}}{r_a} + \frac{GM_0 e^{i\omega(t'')}}{r_b}. \tag{5}$$

Using the fact that gravitational effects propagate at the speed of light c from General Relativity [3], equation (5) can be written as

$$\Phi(\bar{r},t)=\frac{GM_0e^{i\omega(t-\frac{r_a}{c})}}{r_a}+\frac{GM_0e^{i\omega(t-\frac{r_b}{c})}}{r_b}. \tag{6}$$

From figure 1 and the cosine rule it can be shown that

$$r_a\approx r-\frac{s}{2}\cos\theta=r\left(1-\frac{s}{2r}\cos\theta\right)$$

and

$$r_b\approx r+\frac{s}{2}\cos\theta=r\left(1+\frac{s}{2r}\cos\theta\right)$$

and assuming that $r\gg s$ then

$$t-\frac{r_a}{c}=t-\frac{r}{c}+\frac{s}{2c}\cos\theta \tag{7}$$

$$t-\frac{r_b}{c}=t-\frac{r}{c}-\frac{s}{2c}\cos\theta. \tag{8}$$

Substituting equations (7) and (8) into equation (6) yields

$$\Phi(\bar{r},t)=\frac{GM_0}{r}e^{i\omega(t-\frac{r}{c})}\left(\frac{e^{\frac{is}{2\lambda\!\!\!^{-}}\cos\theta}}{1-\frac{s}{2r}\cos\theta}+\frac{e^{-\frac{is}{2\lambda\!\!\!^{-}}\cos\theta}}{1-\frac{s}{2r}\cos\theta}\right) \tag{9}$$

where $\lambda\!\!\!^{-}=\frac{\lambda}{2\pi}=\frac{c}{\omega}$. λ is the wavelength of the gravitational wave.

Series expansion of the exponential term and denominator of the fractions in the brackets of equation (9) to the first power of $\frac{s}{\lambda\!\!\!^{-}}$ and $\frac{s}{r}$ yields

$$\Phi(\bar{r},t)=\frac{2GM_0}{r\lambda\!\!\!^{-}}e^{i\omega(t-\frac{r}{c})}\left(\lambda\!\!\!^{-}+\frac{is^2}{4r}\cos^2\theta\right). \tag{10}$$

Equation (10) is valid provided $s<<r$ and $s<<\lambda$ for arbitrary s and λ.

But from complex analysis it can be shown that,

$$\lambda\!\!\!^{-}+\frac{is^2}{4r}\cos^2\theta=\left(\lambda\!\!\!^{-2}+\frac{s^4\cos^4\theta}{16r^2}\right)^{1/2}e^{i\sigma} \tag{11}$$

where

$$\sigma=\arctan\left(\frac{s^4\cos^4\theta}{16r^2\lambda\!\!\!^{-2}}\right).$$

Thus equation (10) becomes,

$$\Phi(\bar{r},t)=\frac{2GM_0}{r\lambda\!\!\!^{-}}e^{i\omega(t-\frac{r}{c})}\left(\lambda\!\!\!^{-2}+\frac{s^4\cos^4\theta}{16r^2}\right)^{1/2}e^{i\sigma} \tag{12}$$

or

$$\Phi(\bar{r},t)=\frac{2GM_0}{r\lambda\!\!\!^{-}}\left(\lambda\!\!\!^{-2}+\frac{s^4\cos^4\theta}{16r^2}\right)^{1/2}e^{i\omega\left(t-\frac{r}{c}+\frac{1}{\omega}\sigma\right)}. \tag{13}$$

From equation (13), it is deduced that the gravitational potential propagates as a wave with phase $t-\frac{r}{c}+\frac{1}{\omega}\sigma$.

The following remarks can be deduced from the expression of gravitational potential in this field:

- For
$$s^4\cos^4\theta\gg 16r^2\lambda\!\!\!^{-2}$$
it is clear that
$$\arctan\left(\frac{s^4\cos^4\theta}{16r^2\lambda\!\!\!^{-2}}\right)\approx\frac{\pi}{2}.$$
Thus in this case, the phase velocity of the gravitational potential is c.
- If $s^4\cos^4\theta$ is not much greater than $16r^2\lambda\!\!\!^{-2}$ then the phase velocity of propagation is larger than c. This provides a crucial condition for the propagation of gravitational waves from an electric dipole antenna at velocities greater than the speed of light.
- At the equatorial plane of the electric dipole antenna, $\theta=\frac{\pi}{2}$ and
$$\Phi(\bar{r},t)=\frac{2GM_0}{r}e^{i\omega(t-\frac{r}{c})}.$$
This indicates that at the equatorial plane; the gravitational wave propagates at a phase velocity of c, unlike in the case of electromagnetic waves, where fields of the two electrodes cancel out each other due to charge symmetry.
- Also, the gravitational field varies as $\frac{1}{r}$ and thus the wave dies out as one moves away from the dipole antenna. This is in agreement with the prediction by Astrophysicists that as gravitational waves travel from galaxies towards the Earth, their intensities die off and they become too weak when they get to planet Earth.

3 Conclusion

The major significance of this article is that, although the electric dipole antenna is not made up of massive compact bodies, the generation of gravitational radiation has been shown theoretically. Hence, this article highlights the fact that gravitational radiation can be produced by an interaction of two masses irrespective of their sizes. The use of gravitational potential which is a dynamical parameter also signifies that the existence of gravitational waves can also be predicted using Newton's theory of gravitation.

Submitted on February 07, 2013 / Accepted on February 27, 2013

References

1. Jones D.S. The Theory of Electromagnetism. Macmillan, New York, 1964.
2. Howusu S.X.K. Complete Dynamical Theories of Physics. Jos University Press, 2010.
3. Weinberg S. Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity. J. Wiley, New York, 1972.

The Heisenberg Uncertainty Principle and the Nyquist-Shannon Sampling Theorem

Pierre A. Millette

E-mail: PierreAMillette@alumni.uottawa.ca, Ottawa, Canada

The derivation of the Heisenberg Uncertainty Principle (HUP) from the Uncertainty Theorem of Fourier Transform theory demonstrates that the HUP arises from the dependency of momentum on wave number that exists at the quantum level. It also establishes that the HUP is purely a relationship between the effective widths of Fourier transform pairs of variables (i.e. conjugate variables). We note that the HUP is not a quantum mechanical measurement principle *per se*. We introduce the Quantum Mechanical equivalent of the Nyquist-Shannon Sampling Theorem of Fourier Transform theory, and show that it is a better principle to describe the measurement limitations of Quantum Mechanics. We show that Brillouin zones in Solid State Physics are a manifestation of the Nyquist-Shannon Sampling Theorem at the quantum level. By comparison with other fields where Fourier Transform theory is used, we propose that we need to discern between measurement limitations and inherent limitations when interpreting the impact of the HUP on the nature of the quantum level. We further propose that while measurement limitations result in our perception of indeterminism at the quantum level, there is no evidence that there are any inherent limitations at the quantum level, based on the Nyquist-Shannon Sampling Theorem.

1 Introduction

The Heisenberg Uncertainty Principle is a cornerstone of quantum mechanics. As noted by Hughes [1, see pp. 265–266], the interpretation of the Principle varies

- from expressing a limitation on measurement as originally derived by Heisenberg [2] (Heisenberg's microscope),
- to being the variance of a measurement carried out on an ensemble of particles [3] [4],
- to being inherent to a microsystem [5], meaning essentially that there is an indeterminism to the natural world which is a basic characteristic of the quantum level.

Greenstein retains only the first and last alternatives [6, see p. 51].

However, the Heisenberg Uncertainty Principle can be derived from considerations which clearly demonstate that these interpretations of the principle are not required by its mathematical formulation. This derivation, based on the application of Fourier methods, is given in various mathematical and engineering textbooks, for example [7, see p. 141].

2 Consistent derivation of the Heisenberg Uncertainty Principle

In the Fourier transform literature, the Heisenberg Uncertainty Principle is derived from a general theorem of Fourier theory called the Uncertainty Theorem [7]. This theorem states that the effective width of a function times the effective width of its transform cannot be less than a minimum value given by

$$W(f)\,W(\tilde{f}) \geqslant 1/2 \tag{1}$$

where f is the function of interest and $\tilde{f}$ is its Fourier transform. $W(f)$ is the effective width of function f, defined by

$$|W(f)|^2 = \frac{\int_{-\infty}^{\infty} |f(u)|^2 [u - M(f)]^2 du}{\int_{-\infty}^{\infty} |f(u)|^2 du} \tag{2}$$

and $M(f)$ is the mean ordinate defined by

$$M(f) = \frac{\int_{-\infty}^{\infty} |f(u)|^2 u du}{\int_{-\infty}^{\infty} |f(u)|^2 du}. \tag{3}$$

There are several points that must be noted with respect to this derivation:

Eq.(1) applies to a Fourier transform pair of variables. Taking the simple case of time t and frequency ν to illustrate the point: If we consider the function f to be the function that describes a time function t, then the width of the function, $W(f)$, can be denoted as $W(f) = \Delta t$. The Fourier transform of function t is the frequency function ν and the width of this function can be denoted as $W(\tilde{t}) = W(\nu) = \Delta\nu$. Substituting in (1), the Uncertainty Theorem then yields

$$\Delta t\,\Delta\nu \geqslant 1/2. \tag{4}$$

However, if one wishes to use the circular frequency $\omega = 2\pi\nu$ instead, (4) becomes

$$\Delta t\,\Delta\omega \geqslant \pi. \tag{5}$$

It is thus necessary to take special care to clearly identify the Fourier transform variable used as it impacts the R.H.S. of the resulting Uncertainty relation (see for example [8] and [9, pp. 21–22]).

Equations (4) and (5) above correspond to the following definitions of the Fourier transform respectively [8]:

Equation (4):

$$f(t) = \int_{-\infty}^{\infty} \tilde{f}(\nu) \exp(2\pi i \nu t) d\nu \tag{6}$$

$$\tilde{f}(\nu) = \int_{-\infty}^{\infty} f(t) \exp(-2\pi i \nu t) dt \tag{7}$$

Equation (5):

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \tilde{f}(\omega) \exp(i\omega t) d\omega \tag{8}$$

$$\tilde{f}(\omega) = \int_{-\infty}^{\infty} f(t) \exp(-i\omega t) dt \tag{9}$$

Sometimes the factor $1/2\pi$ is distributed between the two integrals (the Fourier and the Inverse Fourier Transform Integrals) as $1/\sqrt{2\pi}$. In Physics, (8) and (9) are preferred, as this eliminates the cumbersome factor of 2π in the exponential (see for example [10, p. 12]), but care must then be taken to ensure the resulting factor of $1/2\pi$ in (8) is propagated forward in derivations using that definition.

Using the relation $E = h\nu$, where h is Planck's constant, in (4) above, or the relation $E = \hbar\omega$, where $\hbar = h/2\pi$, in (5) above, one obtains the same statement of the Heisenberg Uncertainty Principle namely

$$\Delta E \, \Delta t \geqslant h/2 \tag{10}$$

in both cases.

Similarly for the position x, if we consider the function f to be the function that describes the position x of a particle, then the width of the function, $W(f)$, can be denoted as $W(f) = \Delta x$. The Fourier transform of function x is the function $\tilde{x} = \lambda^{-1}$ and the width of this function can be denoted as $W(\tilde{x}) = W(\lambda^{-1}) = \Delta(\lambda^{-1})$ which we write as $\Delta\lambda^{-1}$ for brevity. You will note that we have not used the wavenumber function k, as this is usually defined as $k = 2\pi/\lambda$ (see for example [11] and references). Substituting in (1), we obtain the relation

$$\Delta x \, \Delta\lambda^{-1} \geqslant 1/2. \tag{11}$$

In terms of the wavenumber k, (11) becomes

$$\Delta x \, \Delta k \geqslant \pi. \tag{12}$$

Given that the momentum of a quantum particle is given by $p = h/\lambda$ or by $p = \hbar k$, both (11) and (12) can be expressed as

$$\Delta x \, \Delta p \geqslant h/2. \tag{13}$$

Equations (10) and (13) are both different statements of the Heisenberg Uncertainty Principle.

The R.H.S. of these equations is different from the usual statement of the Heisenberg Uncertainty Principle where the value $\hbar/2$ is used instead of the value $h/2$ obtained in this analysis. The application of (4) to circular variables (i.e. using ω in (4) instead of (5)) would result in the (incorrect) expression

$$\Delta t \, \Delta\omega \geqslant 1/2 \tag{14}$$

and the more commonly encountered (incorrect) expression

$$\Delta E \, \Delta t \geqslant \hbar/2. \tag{15}$$

However, Heisenberg's original derivation [2] had the R. H.S. of (13) approximately equal to h, and Greenstein's rederivation [6, see p. 47] of Heisenberg's principle results in the value $h/2$. Kennard's formal derivation [12] using standard deviations established the value of $\hbar/2$ used today. This would thus seem to be the reason for the use of the value $\hbar/2$ in the formulation of the Heisenberg Uncertainty Principle.

Recently, Schürmann et al [13] have shown that in the case of a single slit diffraction experiment, the standard deviation of the momentum typically does not exist. They derive the conditions under which the standard deviation of the momentum is finite, and show that the R.H.S. of the resulting inequality satisfies (13). It thus seems that (13) is the more general formulation of the Heisenberg Uncertainty Principle, while the expression with the value $\hbar/2$ derived using standard deviations is a more specific case.

Whether one uses $\hbar/2$ or $h/2$ has little impact on the Heisenberg Uncertainty Principle as the R.H.S. is used to provide an order of magnitude estimate of the effect considered. However, the difference becomes evident when we apply our results to the Brillouin zone formulation of Solid State Physics (as will be seen in Section 5) since this now impacts calculations resulting from models that can be compared with experimental values.

3 Interpretation of the Heisenberg Uncertainty Principle

This derivation demonstrates that the Heisenberg Uncertainty Principle arises because x and p form a Fourier transform pair of variables. It is a characteristic of Quantum Mechanics that conjugate variables are Fourier transform pairs of variables. Thus the Heisenberg Uncertainty Principle arises because the momentum p of a quantum particle is proportional to the de Broglie wave number k of the particle. If momentum was not proportional to wave number, the Heisenberg Uncertainty Principle would not exist for those variables.

This argument elucidates why the Heisenberg Uncertainty Principle exists. Can it shed light on the meaning of the Heisenberg Uncertainty Principle in relation to the basic nature of the quantum level? First, we note that the Uncertainty Principle, according to Fourier transform theory, relates the effective width of Fourier transform pairs of functions or variables. It is not a measurement theorem *per se*. It does not describe what happens when Fourier transform variables are

measured, only that their effective widths must satisfy the Uncertainty Principle.

Indeed, as pointed out by Omnès [14, see p. 57], "it is quite legitimate to write down an eigenstate of energy at a well-defined time". Omnès ascribes this seeming violation of the Heisenberg Uncertainty Principle to the fact that time is not an observable obtained from an operator like momentum, but rather a parameter. Greenstein [6, see p. 65] makes the same argument. However, time t multiplied by the speed of light c is a component of the 4-vector x^μ and energy E divided by c is a component of the energy-momentum 4-vector P^μ. The time component of these 4-vectors should not be treated differently than the space component. The operator versus parameter argument is weak.

What Omnès' example shows is that the impact of the effective widths Δt and ΔE of the Heisenberg Uncertainty Principle depends on the observation of the time function t and of the energy function E that is performed. A time interval Δt can be associated with the time function t during which is measured the energy eigenstate function E which itself has a certain width ΔE, with both widths (Δ) satisfying (10). This example demonstrates that the Heisenberg Uncertainty Principle is not a measurement theorem as often used. Rather, it is a relationship between the effective widths of Fourier transform pairs of variables that can have an impact on the observation of those variables.

A more stringent scenario for the impact of the energy-time Heisenberg Uncertainty Principle is one where the time and energy functions are small quantities. For example, we consider the impact of Δt on the observation of τ_n, the lifetime of an atom in energy eigenstate n, and the impact of ΔE on the transition energy E_{mn}, for a transition between states n and m during spectral line emission. The conditions to be able to observe τ_n and E_{mn} are:

$$\tau_n \geqslant \Delta t \tag{16}$$

$$E_{mn} \geqslant \Delta E. \tag{17}$$

Using (10) in (16),

$$\tau_n \geqslant \Delta t \geqslant h/(2\Delta E). \tag{18}$$

Hence

$$\Delta E \geqslant \frac{h}{2}\frac{1}{\tau_n}. \tag{19}$$

As state n increases, the lifetime τ_n decreases. Eq.(19) is thus more constrained in the limit of large n. Using the following hydrogenic asymptotic expression for τ_n from Millette et al [15]

$$\tau_n \sim \frac{n^5}{\ln(n)} \tag{20}$$

into (19), (17) becomes

$$E_{mn} \geqslant \Delta E \gtrsim \frac{h}{2}\, k\, \frac{\ln(n)}{n^5} \tag{21}$$

where $1/k$ is the constant of proportionality of (20) given by

$$k = \frac{2^6}{3}\sqrt{\frac{\pi}{3}}\, Z^2\alpha^3 c R_H \tag{22}$$

where Z is the nuclear charge of the hydrogenic ion, α is the fine-structure constant, and R_H is the hydrogen Rydberg constant. Eliminating the middle term, (21) becomes

$$E_{mn} \gtrsim \frac{h}{2}\, k\, \frac{\ln(n)}{n^5}. \tag{23}$$

Applying L'Hôpital's rule, the R.H.S. of the above equation is of order

$$\text{R.H.S.} \sim O\left(\frac{1}{n^5}\right) \text{ as } n \to \infty \tag{24}$$

while the L.H.S. is of order [16, see p. 9]

$$\text{L.H.S.} \sim O\left(\frac{1}{n^2}\right) \text{ as } n \to \infty. \tag{25}$$

Given that (24) tends to zero faster than (25), (23) is satisfied. Both τ_n, the lifetime of the atom in energy eigenstate n, and the transition energy E_{mn} for the transition between states n and m satisfy the conditions for observation of the spectral line emission. Thus for the time interval Δt, given by (16), associated with the time function τ_n for the transition energy function E_{mn} which itself has a certain width ΔE, given by (17), both Δ's satisfy (10) as expected, given the observation of spectral line emission.

4 Quantum measurements and the Nyquist-Shannon Sampling Theorem

At the quantum level, one must interact to some degree with a quantum system to perform a measurement. When describing the action of measurements of Fourier transform variables, one can consider two limiting measurement cases: 1) truncation of the variable time series as a result of a fully interacting measurement or 2) sampling of the variable time series at intervals which we consider to be regular in this analysis, in the case of minimally interacting measurements. As we will see, the action of sampling allows for measurements that otherwise would not be possible in the case of a single minimal interaction.

It should be noted that the intermediate case of a partial measurement interaction resulting for example in a transfer of energy or momentum to a particle can be considered as the truncation of the original time series and the initiation of a new time series after the interaction. The advantage of decomposing measurement actions in this fashion is that their impact on Fourier transform variables can be described by the Nyquist-Shannon Sampling Theorem of Fourier transform theory. This theorem is a measurement theorem for Fourier transform variables based on sampling and truncation operations.

The Nyquist-Shannon Sampling Theorem is fundamental to the field of information theory, and is well known in digital signal processing and remote sensing [17]. In its most basic form, the theorem states that the rate of sampling of a signal (or variable) f_s must be greater than or equal to the Nyquist sampling rate f_S to avoid loss of information in the sampled signal, where the Nyquist sampling rate is equal to twice that of the highest frequency component, f_{max}, present in the signal:

$$f_s \geqslant f_S = 2 f_{max}. \tag{26}$$

If the sampling rate is less than that of (26), aliasing occurs, which results in a loss of information.

In general, natural signals are not infinite in duration and, during measurement, sampling is also accompanied by truncation of the signal. There is thus loss of information during a typical measurement process. The Nyquist-Shannon Sampling theorem elucidates the relationship between the process of sampling and truncating a variable and the effect this action has on its Fourier transform [18, see p. 83]. In effect, it explains what happens to the information content of a variable when its conjugate is measured.

Sampling a variable x at a rate δx will result in the measurement of its conjugate variable $\tilde{x}$ being limited to its maximum Nyquist range value $\tilde{x}_N$ as given by the Nyquist-Shannon Sampling theorem:

$$\tilde{x} \leqslant \tilde{x}_N \tag{27}$$

where

$$\tilde{x}_N = 1/(2\delta x). \tag{28}$$

Combining these two equations, we get the relation

$$\tilde{x}\,\delta x \leqslant 1/2, \quad \text{for } \tilde{x} \leqslant \tilde{x}_N. \tag{29}$$

Conversely, truncating a variable x at a maximum value x_N ($x \leqslant x_N$) will result in its conjugate variable $\tilde{x}$ being sampled at a rate $\delta\tilde{x}$ given by the Nyquist-Shannon Sampling theorem $\delta\tilde{x} = 1/(2x_N)$ resulting in the relation

$$\delta\tilde{x}\, x \leqslant 1/2, \quad \text{for } x \leqslant x_N. \tag{30}$$

The impact of the Nyquist-Shannon Sampling theorem is now considered for a particle's position x and momentum p. Applying the theorem to the case where a particle's trajectory is truncated to x_N, we can write from (30), for $x \leqslant x_N$,

$$x\,\delta\lambda^{-1} \leqslant 1/2, \quad \text{for } x \leqslant x_N \tag{31}$$

or

$$x\,\delta k \leqslant \pi, \quad \text{for } x \leqslant x_N \tag{32}$$

which becomes

$$x\,\delta p \leqslant h/2, \quad \text{for } x \leqslant x_N \tag{33}$$

where δp is the p-domain sampling rate and the x values can be measured up to x_N (corresponding to the equality in the equations above).

Conversely, applying the theorem to the case where a particle's trajectory is sampled at a rate δx, one can also write from (29), for $\tilde{x} \leqslant \tilde{x}_N$, where $\tilde{x}$ stands for either of λ^{-1}, k, or p,

$$\delta x\,\lambda^{-1} \leqslant 1/2, \quad \text{for } \lambda^{-1} \leqslant \lambda_N^{-1} \tag{34}$$

or

$$\delta x\, k \leqslant \pi, \quad \text{for } k \leqslant k_N \tag{35}$$

which becomes

$$\delta x\, p \leqslant h/2, \quad \text{for } p \leqslant p_N \tag{36}$$

where δx is the x-domain sampling rate and k_N is the wave number range that can be measured. For the case where the equality holds, we have $k_N = \pi/\delta x$ where k_N is the Nyquist wave number, the maximum wave number that can be measured with a δx sampling interval.

Sampling in one domain leads to truncation in the other. Sampling (δx) and truncation (x_N) in one domain leads to truncation (k_N) and sampling (δk) respectively in the other. As x and k form a Fourier transform pair in quantum mechanics, the Nyquist-Shannon Sampling theorem must also apply to this pair of conjugate variables. Similar relations can be derived for the E and ν pair of conjugate variables.

5 Implications of the Nyquist-Shannon Sampling Theorem at the quantum level

Equations (32) and (35) lead to the following measurement behaviors at the quantum level:

Lower-bound limit: If the position of a particle is measured over an interval x_N, its wave number cannot be resolved with a resolution better than sampling rate δk as given by (32) with $x = x_N$. If the momentum of a particle is measured over an interval k_N, its position cannot be resolved with a resolution better than sampling rate δx as given by (35) with $k = k_N$.

Upper-bound limit: If the position of a particle is sampled at a rate δx, wave numbers up to k_N can be resolved, while wave numbers larger than k_N cannot be resolved as given by (35). If the momentum of a particle is sampled at a rate δk, lengths up to x_N can be resolved, while lengths longer than x_N cannot be resolved as given by (32).

The lower-bound limit is similar to how the Heisenberg Uncertainty Principle is usually expressed when it is used as a measurement principle, although it is not strictly equivalent. The Nyquist-Shannon Sampling Theorem provides the proper formulation and limitations of this type of measurement.

The upper-bound limit suggests a different type of quantum measurement: regular sampling of a particle's position or momentum. In this case, one can obtain as accurate a measurement of the Fourier transform variable as desired, up to

the Nyquist-Shannon Sampling limit of $h/2$ (i.e. in the interval $[0, h/2]$).

An example of this phenomenon occurs in Solid State Physics where the translational symmetry of atoms in a solid resulting from the regular lattice spacing, is equivalent to an effective sampling of the atoms of the solid and gives rise to the Brillouin zone for which the valid values of k are governed by (35). Setting $\delta x = a$, the lattice spacing, and extending by symmetry the k values to include the symmetric negative values, one obtains [19, see p. 34], [20, see p. 100], [10, see p. 21]:

$$-\pi/a \leqslant k \leqslant \pi/a \tag{37}$$

or alternatively

$$k \leqslant |\pi/a|. \tag{38}$$

This is called the reduced zone scheme and π/a is called the Brillouin zone boundary [21, see p. 307]. The Brillouin zones of Solid State Physics are thus a manifestation of the Nyquist-Shannon Sampling theorem at the quantum level.

In essence, this is a theory of measurement for variables that are Fourier transform pairs. The resolution of our measurements is governed by limitations that arise from the Nyquist-Shannon Sampling theorem. Equations (32) and (35) are recognized as measurement relationships for quantum-mechanical conjugate variables. Currently, Quantum Mechanics only considers the Uncertainty Theorem but not the Sampling Theorem. The two theorems are applicable to Quantum Mechanics and have different interpretations: the Uncertainty Theorem defines a relationship between the widths of conjugate variables, while the Sampling Theorem establishes sampling and truncation measurement relationships for conjugate variables.

The value δx is a sampled measurement and as a result can resolve values of p up to its Nyquist value p_N given by the Nyquist-Shannon Sampling theorem, (36). This is a surprising result as the momentum can be resolved up to its Nyquist value, in apparent contradiction to the Heisenberg Uncertainty Principle. Yet this result is known to be correct as demonstrated by the Brillouin zones formulation of Solid State Physics. Physically this result can be understood from the sampling measurement operation which builds up the momentum information during the sampling process, up to the Nyquist limit p_N. It must be remembered that the Nyquist limit depends on the sampling rate δx as per the Nyquist-Shannon Sampling theorem, (36). The Nyquist value must also satisfy (26) to avoid loss of information in the sampling process, due to aliasing.

This improved understanding of the Heisenberg Uncertainty Principle and its sampling counterpart allows us to clarify its interpretation. This is based on our understanding of the behavior of the Uncertainty Theorem and the Nyquist-Shannon Sampling Theorem in other applications such as, for example, Digital Signal Processing.

6 Measurement limitations and inherent limitations

It is important to differentiate between the measurement limitations that arise from the properties of Fourier transform pairs previously considered, and any inherent limitations that may or may not exist for those same variables independently of the measurement process. Quantum theory currently assumes that the inherent limitations are the same as the measurement limitations. This assumption needs to be re-examined based on the improved understanding obtained from the effect of the Uncertainty and Sampling Theorems in other applications.

The properties of Fourier transform pairs considered in the previous sections do not mean that the underlying quantities we are measuring are inherently limited by our measurement limitations. On the contrary, we know from experience in other applications that our measurement limitations do not represent an inherent limitation on the measured quantities in Fourier Transform theory: for example, in Digital Signal Processing, a signal is continuous even though our measurement of the signal results in discrete and aliased values of limited resolution subject to the Nyquist-Shannon Sampling Theorem (analog and digital representation of the signal). The effective width of the signal and its transform are related by the Uncertainty theorem. Even though the time and frequency evolution of a signal that we measure is limited by our measurement limitations, the time domain and frequency domain signals are both continuous, independently of how we measure them.

The measurement limitations apply equally to the macroscopic level and to the quantum level as they are derived from the properties of Fourier transform pairs of variables which are the same at all scales. However, at the quantum level, contrary to our macroscopic environment, we cannot perceive the underlying quantities other than by instrumented measurements. Hence during a measurement process, the quantum level is limited by our measurement limitations. However, assuming that these measurement limitations represent inherent limitations and form a basic characteristic of the quantum level is an assumption that is not justified based on the preceding considerations. Indeed, the Nyquist-Shannon Sampling Theorem of Fourier Transform theory shows that the range of values of variables below the Heisenberg Uncertainty Principle value of $h/2$ is accessible under sampling measurement conditions, as demonstrated by the Brillouin zones formulation of Solid State Physics.

7 Overlap of the Heisenberg Uncertainty Principle and the Nyquist-Shannon Sampling Theorem

Brillouin zone analysis in Solid State Physics demonstrates that one can arbitrarily measure k from 0 up to its Nyquist limit, as long as the variable x is sampled at a constant rate (rather than performing a single x measurement). The Nyquist-Shannon Sampling Theorem can thus be considered to

cover the range that the Heisenberg Uncertainty Principle excludes.

However, one should recognize that the coverage results from two disparate theorems, and one should be careful not to try to tie the two Theorems at their value of overlap π. The reason is that one expression involves the widths of conjugate variables as determined by (1) to (3), while the other involves sampling a variable and truncating its conjugate, or vice versa as determined by (32) and (35). The equations are not continuous at the point of overlap π. Indeed, any relation obtained would apply only at the overlap π and would have no applicability or physical validity on either side of the overlap.

8 Discussion and conclusion

In this paper, we have shown that a consistent application of Fourier Transform theory to the derivation of the Heisenberg Uncertainty Principle requires that the R.H.S. of the Heisenberg inequality be $h/2$, not $\hbar/2$. This is confirmed when extending the analysis to the Brillouin zones formulation of Solid State Physics.

We have noted that the Heisenberg Uncertainty Principle, obtained from the Uncertainty Theorem of Fourier Transform theory, arises because of the dependency of momentum on wave number that exists at the quantum level. Quantum mechanical conjugate variables are Fourier Transform pairs of variables.

We have shown from Fourier Transform theory that the Nyquist-Shannon Sampling Theorem affects the nature of-measurements of quantum mechanical conjugate variables. We have shown that Brillouin zones in Solid State Physics are a manifestation of the Nyquist-Shannon Sampling Theorem at the quantum level.

We have noted that both the Sampling Theorem and the Uncertainty Theorem are required to fully describe quantum mechanical conjugate variables. The Nyquist-Shannon Sampling Theorem complements the Heisenberg Uncertainty Principle. The overlap of these Theorems at the $h/2$ equality value is a mathematical artifact and has no physical significance.

We have noted that the Uncertainty Theorem and the Nyquist-Shannon Sampling Theorem apply to Fourier Transform pairs of variables independently of the level at which the theorems are applied (macroscopic or microscopic). Conjugate variable measurement limitations due to these Theorems affect how we perceive quantum level events as these can only be perceived by instrumented measurements at that level. However, based on our analysis, quantum measurement limitations affect our perception of the quantum environment only, and are not inherent limitations of the quantum level, as demonstrated by the Brillouin zones formulation of Solid State Physics.

The application of the Nyquist-Shannon Sampling Theorem to the quantum level offers the possibility of investigating new experimental conditions beyond the Brillouin zone example from Solid State Physics considered in this paper, allowing a unique vista into a range of variable values previously considered unreachable due to the Heisenberg Uncertainty Principle. Regular sampling of position allows us to determine momentum below its Nyquist limit, and similarly the regular sampling of momentum will allow us to determine position below its Nyquist limit.

Submitted on February 21, 2013 / Accepted on March 04, 2013

References

1. Hughes R. I. G. The Structure and Interpretation of Quantum Mechanics. Harvard University Press, Cambridge, 1989.
2. Heisenberg W. Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik. *Zeit. für Phys.*, 1927, v. 43, 172–98; quoted in Hughes R. I. G. The Structure and Interpretation of Quantum Mechanics. Harvard University Press, Cambridge, 1989.
3. Popper K. R. Quantum Theory and the Schism in Physics. Rowan and Littlefield, Totowa, NJ, 1982; quoted in Hughes R. I. G. The Structure and Interpretation of Quantum Mechanics. Harvard University Press, Cambridge, 1989.
4. Margenau H. The Nature of Physical Reality. McGraw Hill, New York, 1950; quoted in Hughes R. I. G. The Structure and Interpretation of Quantum Mechanics. Harvard University Press, Cambridge, 1989.
5. Davies P. C. W. Quantum Mechanics. Routledge, London, 1984; quoted in Hughes R. I. G. The Structure and Interpretation of Quantum Mechanics. Harvard University Press, Cambridge, 1989.
6. Greenstein G. and Zajonc A. G. The Quantum Challenge, Modern Research on the Foundations of Quantum Mechanics. Jones and Bartlett Publishers, Sudbury, Mass., 1997.
7. Cartwright M. Fourier Methods for Mathematicians, Scientists and Engineers. Ellis Norwood Ltd, London, 1990.
8. Weisstein E. MathWorld. http://mathworld.wolfram.com/FourierTransform.html.
9. Griffiths R. B. Consistent Quantum Theory. Cambridge University Press, Cambridge, 2002.
10. Ziman J. M. Principles of the Theory of Solids, 2nd ed. Cambridge University Press, Cambridge, 1979.
11. Weisstein E. World of Physics. http://scienceworld.wolfram.com/physics/Wavenumber.html.
12. Kennard E. H. Zur Quantenmechanik einfacher Bewegungstypen. *Zeit. für Phys.*, 1927, v. 44, 326.
13. Schürmann T. and Hoffman I. A closer look at the uncertainty relation of position and momentum. *Found. of Phys.*, 2009, v. 39, 958–63.
14. Omnès R. The Interpretation of Quantum Mechanics. Princeton University Press, Princeton, New Jersey, 1994.
15. Millette P. A. and Varshni Y. P. New asymptotic expression for the average lifetime of hydrogenic levels. *Can. J. Phys.*, 1979, v. 57, 334–5.
16. Bethe H. A. and Salpeter E. E. Quantum Mechanics of One- and Two-Electron Atoms. Plenum Publishing Corp., New York, NY, 1977.
17. Oppenheim A. V. and Willsky A. S., Signals and Systems. Prentice-Hall Inc., Englewood Cliffs, New Jersey, 1983.
18. Brigham E. O. The Fast Fourier Transform. Prentice-Hall Inc., Englewood Cliffs, New Jersey, 1974.
19. Harrison W. A. Electronic Structure and the Properties of Solids. Dover Publications, New York, 1989.
20. Chaikin P. M. and Lubensky T. C. Principles of Condensed Matter Physics. Cambridge University Press, Camvridge, 1995.
21. Kittel C. Introduction to Solid State Physics. John Wiley and Sons, New York, 1971.

The Cause of the Increased Luminosity Distances of Supernovae Recorded in the Cosmological Data

Emmanuil Manousos

Astrophysics Laboratory , Faculty of Physics, National and Kapodistrian University of Athens, Panepistimiopolis, GR 15783 Zographos, Athens, Greece. E-mail: emanoussos@phys.uoa.gr

The law of selfvariations quantitatively determines a slight increase of the masses and charges as the common cause of quantum and cosmological phenomena. It predicts and explains the totality of the cosmological data. In this article we present the prediction of the law concerning the increased luminosity distances of distant astronomical objects. The prediction we make is in agreement with the cosmological data for the luminosity distances of type Ia supernovae.

1 Introduction

The science of Physics possesses today a plethora of knowledge that allows us to seek the first principles governing physical reality. We can search for a small number of propositions-axioms that could reproduce the totality of our knowledge in Physics. The theory of selfvariations has emerged along this line of reasoning.

We make two hypotheses: The rest masses and electric charges of the material particles increase slightly with the passage of time (selfvariations), and the consequences of this increase propagate in four-dimensional spacetime with a vanishing arc length. Starting from these two hypotheses we conclude that the selfvariations occur in a strictly defined manner. We call the quantitative mathematical determination of the way in which the selfvariations occur, the law of selfvariations.

The law of selfvariations contains an exceptionally large amount of data and information. It is related to the quantum phenomena, the potential fields, and the cosmological data. With the evidence we have in our disposal, and the mathematical calculations we have performed, we can propose the law of selfvariations as the common cause of quantum phenomena and cosmological data. The consequences of the law of selfvariations extend from the microcosm up to the observations we conduct billions of light years away. Equation

$$\left(m_0c^2 + ih\frac{\dot{m}_0}{m_0}\right)^{\cdot} = 0,$$

with unique unkown the rest mass m_0 of particles, both contains as physical information, and justifies, the whole corpus of the current cosmological observational data.

Specifically for the cosmological data, the law of selfvariations predicts and justifies: the redshift of distant astronomical objects and Hubble's law, the cosmic microwave background radiation, the large-scale structures of matter in the Universe, the fact that the Universe is flat, the fact that the total energy-content of the Universe is zero, the fact that the very early Universe went through a phase of ionization, the arrow of time in the macrocosm and its breakdown in the microcosm, the fact that the luminosity distances of distant astronomical objects will always be measured larger than the actual distances. It is this last prediction that we present in the current article.

Since the observations of distant astronomical objects correspond to past time, the rest masses of the material particles in these objects are smaller than the corresponding masses we measure in the laboratory, due to the selfvariations. Therefore, the energy resulting from fusion and fission in distant astronomical objects is less than expected. These distant astronomical objects are fuelled with a smaller than expected amount of energy in order to emit the electromagnetic radiation we observe today from Earth. This fact reduces the luminosity of distant astronomical objects.

In the last decade of the previous century two independent research groups under A.G. Riess and S. Perlmutter, measured the decrease of the luminosity for a large number of type Ia supernovae at great distances. In order to explain the observational data within the framework of the standard cosmological model, the hypothesis of dark energy was introduced.

We have today a large amount of observational data confirming the decrease of luminosity at large distances. All these measurements result in a specific diagram correlating the luminosity-distance with the redshift of distant astronomical objects. This diagram, as it results from the cosmological data, is exactly the same with the one predicted theoretically by the law of selfvariations. In the next paragraph we present the diagram that we theoretically predict.

2 The luminosity distances of distant astronomical objects will always be measured greater than their real distances

The law of selfvariations [1, 2] predicts the relation

$$r = \frac{c}{k}\ln\left(\frac{A}{1-(1+z)(1-A)}\right),$$

between the distance r and the redshift z of distant astronomical objects. For the dimensionless parameter A, it holds that

$A \to 1^-$, since it obeys the inequality

$$\frac{z}{1+z} < A < 1,$$

for every value of the redshift z. The parameter k is constant, and is related to the Hubble parameter H through equation

$$\frac{kA}{1-A} = H.$$

The law of selfvariations predicts that the energy $E(z)$ resulting from fusion, and which powers the distant astronomical objects, is decreased compared with the corresponding energy E measured in the laboratory, according to relation

$$E(z) = \frac{E}{1+z}\,.$$

Because of this, the luminosity of distant astronomical objects is decreased, relative to the expected one. This has as a consequence that the luminosity distances R of distant astronomical objects are measured larger than the actual distances r, $R > r$. From the mathematical calculations [1, 2] we obtain

$$R = r\sqrt{1+z}\,,$$

between the distances R and r.

Combining the previous equations we get the luminosity distance R as a function of the redshift z of distant astronomical objects:

$$R = \frac{cA\sqrt{1+z}}{(1-A)H} \ln\left(\frac{A}{1-(1+z)(1-A)}\right).$$

In the diagram in figure 1 we present the diagram of $R = R(z)$ for $A = 0.975$, $A = 0.990$, $A = 0.995$, $A = 0.999$, $H = 60\,\text{km/s Mpc}$, $c = 3 \times 10^5\,\text{km/s}$ up to $z = 1.5$. In order to explain the inconsistency of the Standard Cosmological Model with the diagram in figure 1, the existence of dark energy was invented and introduced.

Type Ia supernovae are astronomical objects for which we can measure their luminosity distance for great distances. The measurements already conducted [3, 4] agree with the diagram in figure 1.

In the measurements conducted for the determination of the Hubble parameter H, the consequences of equation $R = r\sqrt{1+z}$ have not been taken into account. For small values of the redshift z, the value $H = 60\,\text{km/s Mpc}$ results. The measurements made up to date, have included astronomical objects with a high redshift z, thus raising the value of parameter H to between 72 and 74 km/s Mpc. Today we perform measurements of very high accuracy. Taking into consideration the consequences of equation $R = r\sqrt{1+z}$, we predict that the value of parameter H will be measured independently of the redshift z of the astronomical object. We, of course, refer to measurements of the parameter H that are based on the luminosity distance of the astronomical objects.

Submitted on: March 03, 2013 / Accepted on March 08, 2013

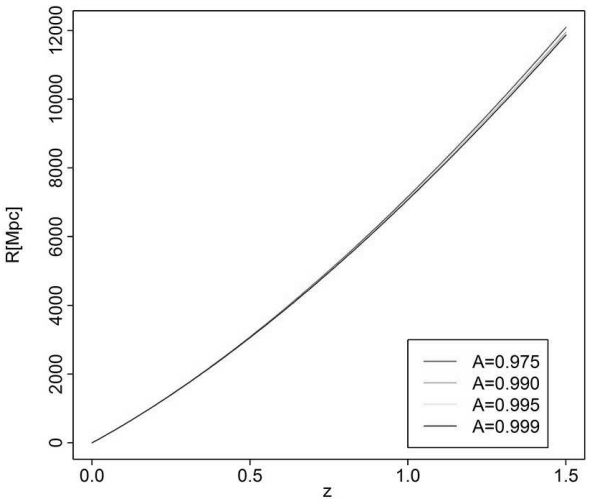

Fig. 1: The diagram of $R = R(z)$ for $A = 0.975$, $A = 0.990$, $A = 0.995$, $A = 0.999$, $H = 60\,\text{km/s Mpc}$, $c = 3 \times 10^5\,\text{km/s}$ up to $z = 1.5$. The measurement of the luminosity distances of type Ia supernova confirms the theoretical prediction of the law of selfvariations.

References

1. Manousos E. A common underlying cause for quantum phenomena and cosmological data. viXra:1302.0115 v2.
2. Manousos E. The theory of self-variations. A continuous slight increase of the charges and the rest masses of the particles can explain the cosmological data. *Nuovo Cimento B*, 2007, 359–388.
3. Adam G. Riess et al. Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant. *The Astronomical Journal*, 1998, v. 29 (3), 423–428.
4. Perlmutter, S. et al. Measurements of Omega and Lambda from 42 High-Redshift Supernovae. *The Astrophysical Journal*, 1999, v. 517 (2), 565–586.

Intrinsic Charges and the Strong Force

Bo Lehnert

Alfvén Laboratory, Royal Institute of Technology, 10044 Stockholm, Sweden. E-mail: Bo.Lehnert@ee.kth.se

According to a revised quantum electrodynamic theory, there are models of leptons such as the electron which possess both a net integrated electric charge and a much larger intrinsic charge of both polarities. From estimates based on such models, the corresponding Coulomb force due to the intrinsic charges then becomes two orders of magnitude larger than that due to the conventional net charge. This intrinsic charge force can also have the features of a short-range interaction. If these results would generally hold true, the intrinsic charge force could either interact with a strong force of different origin and character, or could possibly become identical with the strong force.

1 Introduction

According to quantum mechanics there exists a nonzero lowest energy level, the Zero Point Energy. The vacuum is therefore not merely an empty space, but includes a "photon gas" of related electromagnetic vacuum fluctuations. The pressure of this gas is a physical reality, as demonstrated by the force between two metal plates proposed by Casimir [1] and first confirmed experimentally by Lamoreaux [2].

These circumstances have formed the starting point of a revised quantum electrodynamical approach by the author [3]. In the latter a *nonzero* electric field divergence div **E** is introduced in the *vacuum* state. In its turn, the nonzero electric field divergence admits an additional degree of freedom into the electromagnetic field equations. The latter then possess new solutions both in the steady and the time-dependent states, having applications to modified models of leptons and photons.

In this paper an example is given in Section 2 on the consequences of a nonzero electric field divergence in a steady state. It demonstrates that the local variations of the charge density $\bar{\rho}=\varepsilon_0 \mathrm{div}\,\mathbf{E}$ can result in considerable *intrinsic* charges of both signs, being much larger than the total *net* integrated charge. The possible effects of the intrinsic charges on the Coulomb interaction will then be outlined in Section 3, first in respect to the magnitude of the resulting forces, and then to the range of the same forces in a simple "Gedanken experiment". In Section 4 a comparison is finally made to the strong nuclear force.

2 An Example given by the Revised Electron Model

In the revised quantum electrodynamic theory there are steady states which do not exist in conventional theory [3]. These states include net as well as intrinsic electric charges, electric currents, static electromagnetic fields and related forces. To illustrate the resulting charge distributions, an example is here taken from a corresponding electron model. The features of the model will shortly be summarized here, with reference to details in the original descriptions [3].

In the revised theory the field configuration is shown to become derivable from a generating function

$$F = G_0\, G\,(\rho,\theta) \qquad G = R\,(\rho)\cdot T\,(\theta) \tag{1}$$

in spherical coordinates (r,θ,φ) of an axisymmetric case being independent of the angle φ. Here G_0 stands for a characteristic amplitude, $\rho = r/r_0$ with r_0 as a characteristic radial length, and

$$R = \rho^{-\gamma} e^{-\rho} \qquad \gamma > 0 \tag{2}$$

$$T = 1 + a_1 \sin\theta + a_2 \cos 2\theta + a_3 \sin 3\theta + a_4 \cos 4\theta + \ldots \tag{3}$$

with a_1, a_2, a_3, ... as constant amplitude factors. The radial function R has to be divergent at the origin $r = 0$ to result in a net integrated charge. Thereby a revised renormalisation procedure is applied to make this divergence result in a finite net integrated charge. This leads to forms of the net charge q_0, magnetic moment M_0, rest mass m_0, and angular momentum (spin) s_0 as given by

$$q_0 = 2\pi\varepsilon_0 c_{rG} A_q, \tag{4}$$

$$s_0 = \frac{1}{2}\pi\left(\frac{\varepsilon_0}{c^2}\right) C c_{rG}^2 A_s, \tag{5}$$

$$M_0 m_0 = \left(\frac{\pi\varepsilon_0}{c}\right)^2 C c_{rG}^3 A_M A_m. \tag{6}$$

Here $C = \pm c$, c_{rG} is a finite counter factor in the renormalisation process, and

$$A_k = \int_0^{\pi} I_{k\theta}\, d\theta \qquad k = q, M, m, s \tag{7}$$

with $I_{k\theta}$ being functions of the amplitude factors of equation (3) and the variable $s \equiv \sin\theta$. The factor c_{rG} includes the amplitude G_0 which can have either sign and becomes negative in the case of the electron.

Two quantum conditions are considered here. The first is $s_0 = \pm h/4\pi$ on the spin which results in a normalized net charge

$$q^* = \left|\frac{q_0}{e}\right| = \left(\frac{f_0 A_q^2}{A_s}\right)^{1/2} \qquad f_0 = \frac{2\varepsilon_0 c h}{e^2} \tag{8}$$

where e is the experimentally determined elementary charge and $f_0 \cong 137.036$ is the inverted value of the fine-structure constant. The second condition concerns the magnetic moment and becomes

$$\frac{M_0 m_0}{q_0 s_0} = \frac{A_M A_m}{A_q A_s} = 1 + \delta_M \tag{9}$$

with $\delta_M = 1/2\pi f_0 \cong 0.0011614$. Here it has to be observed that the fourteenth term in equation (7.56) of Reference [3] should read $-699.7897637 a_2 a_3$.

In the four-amplitude case (a_1, a_2, a_3, a_4) the normalized charge q^* will here be studied with conditions (8) and (9) imposed, and as functions of a_3 and a_4 in $a_3 a_4$-space. Then q^* is found to have a minimum for large positive values of a_3 and a_4, within a narrow channel positioned around a plateau defined by the experimental value $q^* = 1$. The width of the channel is only a few percent of q^*. At the plateau the amplitude values are therefore replaced by

$$\bar{a}_i = a_i / a_\infty \qquad a_\infty \gg 1 \qquad i = 1, 2, 3 \ldots \tag{10}$$

As an illustration of the resulting intrinsic and net electric charges, we now use an example where $\bar{a}_1 = -1.91$, $\bar{a}_2 =$ -2.51, $\bar{a}_3 = \bar{a}_4 = 1$. The corresponding integrand $\bar{I}_{q\theta}$ of equation (7) in the plateau region then becomes

$$\begin{aligned} \bar{I}_{q\theta} &= 2sT - 4s^3 T - sD_\theta T + \\ &\quad + 2s^3 D_\theta T + 2sD_\theta\left(s^2 T\right) - sD_\theta\left(s^2 D_\theta T\right) = \\ &= 44.9s + 288s^2 - 2159s^3 - 1320s^4 + \\ &\quad + 7559s^5 + 1120s^6 - 5760s^7 \end{aligned} \tag{11}$$

with the operator

$$D_\theta = -\frac{\partial^2}{\partial\theta^2} - \frac{\cos\theta}{\sin\theta}\frac{\partial}{\partial\theta}. \tag{12}$$

From the corresponding equation (7) this yields $\bar{A}_q \cong 4.600$, $\bar{A}_M \cong 4.37$, $\bar{A}_m \cong 2832$, $\bar{A}_s \cong 2648$ and results in

$$\frac{\bar{A}_M \bar{A}_m}{\bar{A}_q \bar{A}_s} \cong 1.017$$

and $q^* \cong 1.046$.

The obtained value of $\bar{A}_q$ corresponds to the net charge q_0 of equation (4). The detailed charge distribution as a function of s is given by equation (11) and has been plotted in Fig. 1. According to the figure the negative part of the intrinsic charge in the range $0 < \theta < \pi$ is estimated to have the corresponding value $\bar{A}_{q-} \cong 117.3$. The positive part of the intrinsic charge further corresponds to $\bar{A}_{q+} = \bar{A}_{q-} + \bar{A}_q \cong 121.9$.

In the example given here there is thus an outbalanced intrinsic charge proportional to $\bar{A}_{q-} = \bar{A}_{q+} - \bar{A}_q$, plus a net integrated charge proportional to $\bar{A}_q$. The ratio between these charges becomes

$$c_{in} = \frac{\bar{A}_{q-}}{\bar{A}_q}. \tag{13}$$

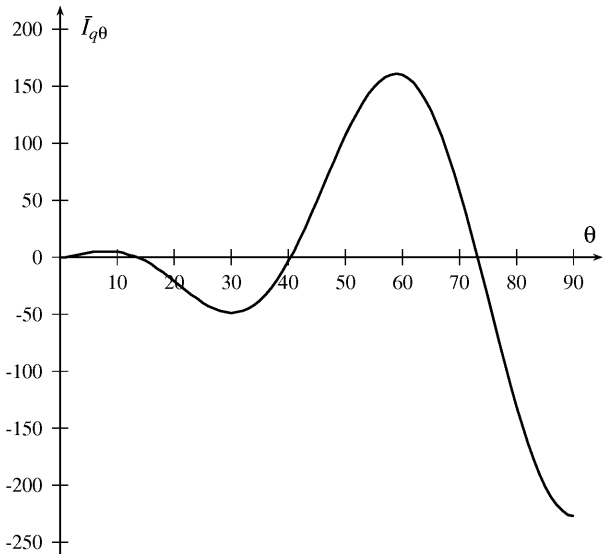

Fig. 1: The local contribution $\bar{I}_{q\theta}$ of equation (10) to the electric charge integral $\bar{A}_q$ as a function of the polar coordinate θ, in the range $0 < \theta < \pi/2$ and with θ given in degrees.

In the example of Fig. 1 it has the value $c_{in} \cong 26.5$, thus indicating that the intrinsic charge considerably exceeds the net charge.

Also the electromagnetic force

$$\mathbf{f} = \bar{\rho}\,(\mathbf{E} + \mathbf{C} \times \mathbf{B}) \tag{14}$$

per unit volume has to be taken into account. It consists of the electrostatic and magnetostatic contributions $\bar{\rho}\mathbf{E}$ and $\bar{\rho}\mathbf{C} \times \mathbf{B}$ where $\mathbf{E}$ and $\mathbf{B}$ are the electric and magnetic field strengths, the velocity vector $\mathbf{C}$ has the modulus $|\mathbf{C}| = \pm c$ and c is the velocity of light. In a cylindrically symmetric case the local electric and magnetic contributions can outbalance each other, but only partly in a spherical axisymmetric case [3]. In the latter case the average radial force can on the other hand be balanced at least for specific solutions of the field equations, but this requires further detailed analysis in every case.

3 Intrinsic Coulomb Forces

The intrinsic charge ratio c_{in} is likely to have consequences when considering the mutual Coulomb forces.

3.1 General Aspects

For any distribution of electric charges the local contribution Δf_{12} to the mutual Coulomb force becomes

$$\Delta f_{12} = \frac{(\Delta q_1)\,(\Delta q_2)}{4\pi\varepsilon_0 r_{12}^2} \tag{15}$$

where and Δq_1 and Δq_2 are two interacting charge elements separated by the distance r_{12}. The charge ratio of equation

(13) thus predicts that the intrinsic Coulomb forces in some cases even may be represented by a factor c_{in}^2 as compared to those in a conventional analysis. For the values of $c_{in}^2 \cong 702$ in the example of Section 2 these forces could then roughly be estimated to be more than two orders of magnitude larger than the conventional ones. However, the effective magnitude of the intrinsic charge force will also depend on the specific geometry of the charge distribution, as being demonstrated by a simple discussion in the following subsection.

3.2 A Gedanken Experiment

To crudely outline the forces which can arise from the intrinsic charges, a simple "Gedanken experiment" is now performed according to Fig. 2. It concerns the interaction between two rigid mutually penetrable spherical configurations, $(_1)$ and $(_2)$, of charge $+Q$ at their centra and charge $-Q$ at their peripheries. The resulting electrostatic field strengths are $\mathbf{E}_1$ and $\mathbf{E}_2$, and the external space is field-free. When these configurations, being simulated as "particles", are apart as in Fig. 2(a), their mutual interaction force F_{12} remains zero. As soon as particle $(_2)$ starts to penetrate particle $(_1)$, part of the negative charge cloud at the periphery of particle $(_2)$ will interact with the electric field $\mathbf{E}_1$ of particle $(_1)$. This generates an attractive force $F_{12} > 0$, as shown by Fig. 2(b). When particle $(_2)$ further penetrates into the field region of particle $(_1)$, however, the mutual interaction force $F_{12} < 0$ changes sign and becomes repulsive as shown in Fig. 2(c). Between cases (a) and (b) there is an equilibrium with $F_{12} = 0$.

The relative magnitude of the maximum force F_{12} in the case of Fig. 2(b) can be estimated by noticing that it is generated by the fraction g_2 of the charge $-Q$ at the periphery of particle $(_2)$, in the field $\mathbf{E}_1$ of particle $(_1)$. With the charge ratio

$$c_{in} = \frac{Q}{e} \tag{16}$$

of the particles $(_1)$ and $(_2)$ this yields an estimated ratio

$$f_{in} = g_2\, c_{in}^2 \tag{17}$$

between the intrinsic forces and those which would have been present in a conventional case. With an estimated factor $g_2 \cong 1/4$ for the fraction of negative charge of particle $(_2)$ being present in the field $\mathbf{E}_1$ of Fig. 2(b), and with $c_{in} \cong 702$ due to the example of Section 2, this results in the force ratio $f_{in} \cong 176$.

In reality, however, the mutual interaction in Fig. 2(b) and Fig. 2(c) becomes more complex and includes a rearrangement of the charge geometry. Thus, even if these simple considerations are somewhat artificial, they appear to indicate that the intrinsic Coulomb forces can become about two orders of magnitude larger than the conventional ones. The intrinsic forces can also in some cases have the character of a short-range interaction.

Provided that the present model of charged leptons also can be applied in a first crude approximation to a bound quark, its characteristic radius r_ε can be estimated. It would become $r_\varepsilon = c_{rG}/c_G$ where c_{rG} and c_G are counter factors of a revised renormalisation procedure [3]. This results in radii in the range $10^{-16} < r_\varepsilon < 10^{-14}$ m for the u, d and s quarks.

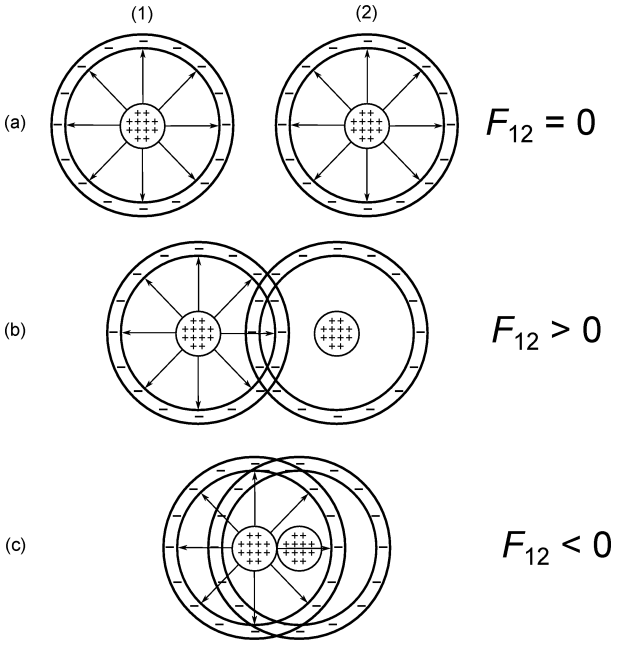

Fig. 2: "Gedanken experiment" where two rigid mutually penetrable spherical configurations ("particles"), $(_1)$ and $(_2)$, are approaching each other. The "particles" have charges $+Q$ at their centra, and $-Q$ at their peripheries, resulting in the internal electric field strengths $\mathbf{E}_1$ and $\mathbf{E}_2$. The mutual interaction force F_{12} is zero when the particles are apart in (a), $F_{12} > 0$ is attractive when they first start to interact in (b), and $F_{12} < 0$ is finally repulsive when they are close together in (c).

4 A Comparison to the Strong Force

The strong force keeps the atomic nucleus together, and it acts on its smallest constituents, the quarks. As concluded from experiments on deep inelastic scattering of energetic electrons by hadrons, the latter include the quarks. According to reviews by French [4], Walker [5] and others, these strong forces have the following features:

- They are primarily attractive.
- They seem to be essentially the same for neutrons and protons.
- Their range is short and not greater than 2×10^{-15} m.
- Within this range they are very strong, i.e. two orders of magnitude larger than those due to conventional electromagnetics.

The strong force can be compared to the intrinsic Coulomb force discussed in this context, also in respect to a possible quark model being somewhat similar to that of the electron as

described in Section 2. The following points should then be noticed:

- The present considerations suggest that the intrinsic charge force can become two orders of magnitude larger than that due to the conventional net charge. The intrinsic charge force thus appears to be of the same order as the strong force, and may also appear in terms of a short-range interaction, on scales of the order of 10^{-15} m.
- It then follows that the intrinsic charge force either will interact with a strong force of different origin and character, or will possibly become identical with the strong force.

Submitted on March 22, 2013 / Accepted on March 22, 2013

References

1. Casimir H. B. G. On the Attraction between two Perfectly Conducting Plates. *Proc. Kon. Nederland Akad. Wetensch.*, 1948, v. B51, 793–795.
2. Lamoreaux S. K. Demonstration of the Casimir Force in the 0.6 to 6 μm Range. *Physical Review Letters*, 1997, v. 78, 5–8.
3. Lehnert B. Revised Quantum Electrodynamics, Nova Science Publishers, New York, 2013, Ch. 6 and Ch. 7.
4. French A. P. Principles of Modern Physics, New York, John Wiley and Sons, 1958, Ch. 9.10.
5. Walker J. S. Physics, Fourth Edition, Pearson International Edition, Addison-Wesley, 2010, page 1144.

Relations between Distorted and Original Angles in STR

Florentin Smarandache
Arts and Science Division, University of New Mexico, 705 Gurley Ave., Gallup, NM 87301, USA. E-mail: smarand@unm.edu

Using the Oblique-Length Contraction Factor, which is a generalization of Lorentz Contraction Factor, one shows several trigonometric relations between distorted and original angles of a moving object lengths in the Special Theory of Relativity.

1 Introduction

The lengths at oblique angle to the motion are contracted with the Oblique-Length Contraction Factor $OC(v,\theta)$, defined as [1-2]:

$$OC(v,\theta) = \sqrt{C(v)^2\cos^2\theta + \sin^2\theta} \tag{1}$$

where $C(v)$ is just Lorentz Factor:

$$C(v) = \sqrt{1-\frac{v^2}{c^2}} \in [0,1] \text{ for } v \in [0,c]. \tag{2}$$

Of course

$$0 \le OC(v,\theta) \le 1. \tag{3}$$

The Oblique-Length Contraction Factor is a generalization of Lorentz Contractor $C(v)$, because: when $\theta = 0$, or the length is moving along the motion direction, then $OC(v,0) = C(v)$. Similarly

$$OC(v,\pi) = OC(v,2\pi) = C(v). \tag{4}$$

Also, if $\theta = \pi/2$, or the length is perpendicular on the motion direction, then $OC(v,\pi/2) = 1$, i.e. no contraction occurs. Similarly $OC(v,\frac{3\pi}{2}) = 1$.

2 Tangential relations between distorted acute angles vs. original acute angles of a right triangle

Let's consider a right triangle with one of its legs along the motion direction (Fig. 1).

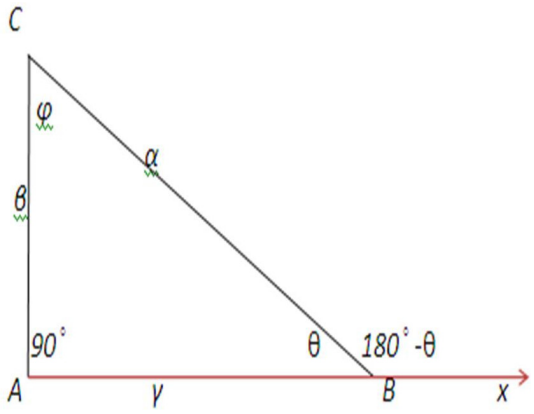

Fig. 1:

$$\tan\theta = \frac{\beta}{\gamma} \tag{5}$$

$$\tan(180^\circ - \theta) = -\tan\theta = -\frac{\beta}{\gamma} \tag{6}$$

After contraction of the side AB (and consequently contraction of the oblique side BC) one gets (Fig. 2):

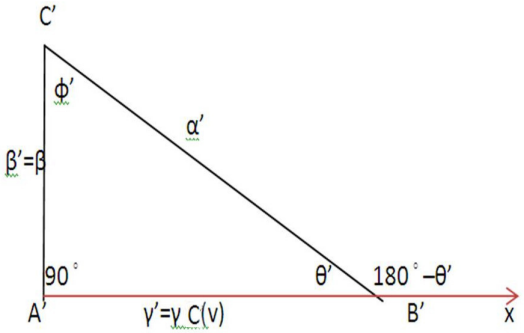

Fig. 2:

$$\tan(180^\circ - \theta') = -\tan\theta' = -\frac{\beta'}{\gamma'} = -\frac{\beta}{\gamma C(v)}. \tag{7}$$

Then:

$$\frac{\tan(180^\circ - \theta')}{\tan(180^\circ - \theta)} = \frac{-\frac{\beta}{\gamma C(v)}}{-\frac{\beta}{\gamma}} = \frac{1}{C(v)}. \tag{8}$$

Therefore

$$\tan(\pi - \theta') = -\frac{\tan(\pi-\theta)}{C(v)} \tag{9}$$

and consequently

$$\tan(\theta') = \frac{\tan(\theta)}{C(v)} \tag{10}$$

or

$$\tan(B') = \frac{\tan(B)}{C(v)} \tag{11}$$

which is the Angle Distortion Equation, where θ is the angle formed by a side travelling along the motion direction and another side which is oblique on the motion direction.

The angle θ is increased (i.e. $\theta' > \theta$).

$$\tan\varphi = \frac{\gamma}{\beta} \quad \text{and} \quad \tan\varphi' = \frac{\gamma'}{\beta'} = \frac{\gamma C(v)}{\beta} \tag{12}$$

whence:

$$\frac{\tan\varphi'}{\tan\varphi} = \frac{\dfrac{\gamma C(v)}{\beta}}{\dfrac{\gamma}{\beta}} = C(v). \tag{13}$$

So we get the following Angle Distortion Equation:

$$\tan\varphi' = \tan\varphi \cdot C(v) \tag{14}$$

or

$$\tan C' = \tan C \cdot C(v) \tag{15}$$

where φ is the angle formed by one side which is perpendicular on the motion direction and the other one is oblique to the motion direction.

The angle φ is decreased (i.e. $\varphi' < \varphi$). If the traveling right triangle is oriented the opposite way (Fig. 3)

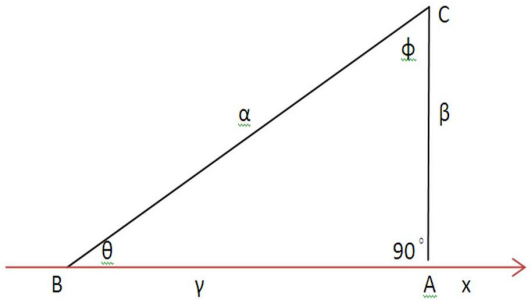

Fig. 3:

$$\tan\theta = \frac{\beta}{\gamma} \quad \text{and} \quad \tan\varphi = \frac{\gamma}{\beta}. \tag{16}$$

Similarly, after contraction of side AB (and consequently contraction of the oblique side BC) one gets (Fig. 4)

$$\tan\theta' = \frac{\beta'}{\gamma'} = \frac{\beta}{\gamma C(v)} \tag{17}$$

and

$$\tan\varphi' = \frac{\gamma'}{\beta'} = \frac{\gamma C(v)}{\beta} \tag{18}$$

$$\frac{\tan\theta'}{\tan\theta} = \frac{\dfrac{\beta}{\gamma C(v)}}{\dfrac{\beta}{\gamma}} = \frac{1}{C(v)} \tag{19}$$

or

$$\tan\theta' = \frac{\tan\theta}{C(v)} \tag{20}$$

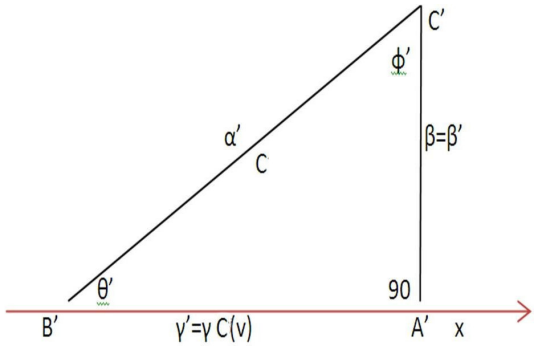

Fig. 4:

and similarly

$$\frac{\tan\varphi'}{\tan\varphi} = \frac{\dfrac{\gamma C(v)}{\beta}}{\dfrac{\gamma}{\beta}} = C(v) \tag{21}$$

or

$$\tan\varphi' = \tan\varphi \cdot C(v). \tag{22}$$

Therefore one got the same Angle Distortion Equations for a right triangle traveling with one of its legs along the motion direction.

3 Tangential relations between distorted angles vs. original angles of a general triangle

Let's suppose a general triangle ΔABC is travelling at speed v along the side BC as in Fig. 5.

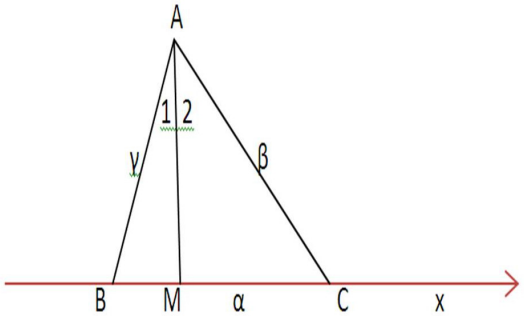

Fig. 5:

The height remains not contracted: $AM \equiv A'M'$. We can split this figure into two traveling right sub-triangles as in Fig. 6.

In the right triangles $\Delta A'M'B'$ and respectively $\Delta A'M'C'$ one has

$$\tan B' = \frac{\tan B}{C(v)} \quad \text{and} \quad \tan C' = \frac{\tan C}{C(v)}. \tag{23}$$

Also

$$\tan A'_1 = \tan A_1 C(v) \quad \text{and} \quad \tan A'_2 = \tan A_2 C(v). \tag{24}$$

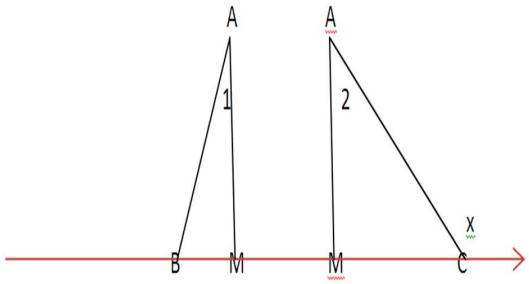

Fig. 6:

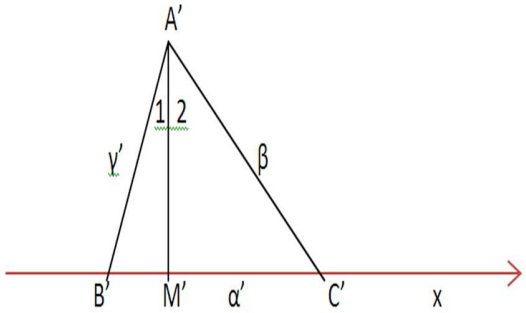

Fig. 7:

But

$$\begin{aligned}\tan A' &= \tan(A'_1 + A'_2) = \frac{\tan A'_1 + \tan A'_2}{1 - \tan A'_1 \tan A'_2} \\ &= \frac{\tan A_1 C(\nu) + \tan A_2 C(\nu)}{1 - \tan A_1 C(\nu) \tan A_2 C(\nu)} \\ &= C(\nu) \cdot \frac{\tan A_1 + \tan A_2}{1 - \tan A_1 \tan A_2 C(\nu)^2} \\ &= C(\nu) \cdot \frac{\dfrac{\tan A_1 + \tan A_2}{1 - \tan A_1 \tan A_2} \cdot (1 - \tan A_1 \tan A_2)}{1 - \tan A_1 \tan A_2 C(\nu)^2} \\ &= C(\nu) \cdot \frac{\tan(A_1 + A_2)}{1} \cdot \frac{1 - \tan A_1 \tan A_2}{1 - \tan A_1 \tan A_2 C(\nu)^2}.\end{aligned}$$

$$\tan A' = C(\nu) \cdot \tan(A) \cdot \frac{1 - \tan A_1 \tan A_2}{1 - \tan A_1 \tan A_2 C(\nu)^2}. \quad (25)$$

We got

$$\tan A' = \tan(A) \cdot C(\nu) \cdot \frac{1 - \tan A_1 \tan A_2}{1 - \tan A_1 \tan A_2 C(\nu)^2} \quad (26)$$

Similarly we can split this Fig. 7 into two traveling right sub-triangles as in Fig. 8.

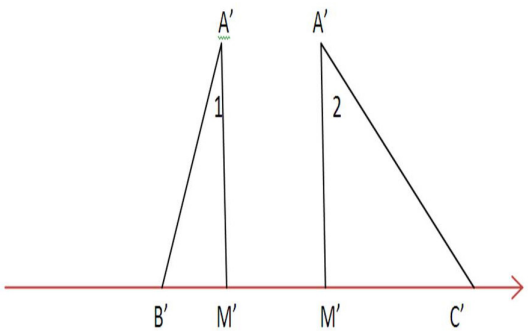

Fig. 8:

4 Other relations between the distorted angles and the original angles

1. Another relation uses the Law of Sine in the triangles ΔABC and respectively $\Delta A'B'C'$:

$$\frac{\alpha}{\sin A} = \frac{\beta}{\sin B} = \frac{\gamma}{\sin C} \quad (27)$$

$$\frac{\alpha'}{\sin A'} = \frac{\beta'}{\sin B'} = \frac{\gamma'}{\sin C'}. \quad (28)$$

After substituting

$$\alpha' = \alpha C(\nu) \quad (29)$$

$$\beta' = \beta \mathscr{C} C(\nu, C) \quad (30)$$

$$\gamma' = \gamma \mathscr{C} C(\nu, B) \quad (31)$$

into the second relation one gets:

$$\frac{\alpha C(\nu)}{\sin A'} = \frac{\beta \mathscr{C} C(\nu, C)}{\sin B'} = \frac{\gamma \mathscr{C} C(\nu, B)}{\sin C'}. \quad (32)$$

Then we divide term by term the previous equalities:

$$\frac{\dfrac{\alpha}{\sin A}}{\dfrac{\alpha C(\nu)}{\sin A'}} = \frac{\dfrac{\beta}{\sin B}}{\dfrac{\beta \mathscr{C} C(\nu, C)}{\sin B'}} = \frac{\dfrac{\gamma}{\sin C}}{\dfrac{\gamma \mathscr{C} C(\nu, B)}{\sin C'}} \quad (33)$$

whence one has:

$$\begin{aligned}\frac{\sin A'}{\sin A \cdot C(\nu)} &= \frac{\sin B'}{\sin B \cdot \mathscr{C} C(\nu, C)} \\ &= \frac{\sin C'}{\sin C \cdot \mathscr{C} C(\nu, B)}.\end{aligned} \quad (34)$$

2. Another way:

$$A' = 180^\circ - (B' + C') \quad \text{and} \quad \mathrm{A} = 180^\circ - (\mathrm{B} + \mathrm{C}) \quad (35)$$

$$\begin{aligned}\tan A' &= \tan[180^\circ - (B' + C')] = -\tan(B' + C') \\ &= -\frac{\tan B' + \tan C'}{1 - \tan B' \cdot \tan C'}\end{aligned}$$

$$
\begin{aligned}
&= -\frac{\dfrac{\tan B}{C(v)} + \dfrac{\tan C}{C(v)}}{1 - \tan B \cdot \tan C/C(v)^2} \\
&= -\frac{1}{C(v)} \cdot \frac{\tan B + \tan C}{1 - \tan B \cdot \tan C/C(v)^2} \\
&= -\frac{\tan(B + C)}{C(v)} \cdot \frac{1 - \tan B \tan C}{1 - \tan B \cdot \tan C/C(v)^2} \\
&= -\frac{-\tan[180^\circ - (B + C)]}{C(v)} \cdot \frac{1 - \tan B \cdot \tan C}{1 - \tan B \cdot \tan C/C(v)^2} \\
&= \frac{\tan A}{C(v)} \cdot \frac{1 - \tan B \cdot \tan C}{1 - \tan B \cdot \tan C/C(v)^2}.
\end{aligned}
$$

We got

$$\tan A' = \frac{\tan A}{C(v)} \cdot \frac{1 - \tan B \cdot \tan C}{1 - \tan B \cdot \tan C/C(v)^2}. \qquad (36)$$

Submitted on March 30, 2013 /Accepted on April 2, 2013

References

1. Smarandache F. New Relativistic Paradoxes and Open Questions. Somipress, Fes, 1983.
2. Smarandache F. Oblique-Length Contraction Factor in the Special Theory of Relativity, *Progress in Physics*, 2013, v. 1, 60–62.
3. Einstein A. On the Electrodynamics of Moving Bodies. *Annalen der Physik*, 1905, v. 17, 891–921.
4. Smarandache F. Absolute Theory of Relativity and Parameterized Special Theory of Relativity and Noninertial Multirelativity. Somipress, Fes, 1982.

The Electron-Vacuum Coupling Force in the Dirac Electron Theory and its Relation to the Zitterbewegung

William C. Daywitt

National Institute for Standards and Technology (retired), Boulder, Colorado. E-mail: wcdaywitt@me.com

From the perspective of the Planck vacuum theory, this paper argues that the standard estimate of the onset radius for electron-positron pair production as the Dirac electron is approached (in its rest frame) is significantly overestimated. The standard value is taken to be the electron Compton radius, while the estimate derived here from the coupling force is over four times smaller. The resulting separation of the Compton radius from the onset radius leads to a clear explanation of the zitterbewegung in terms of vacuum dynamics, making the zitterbewegung a relevant part of the electron theory.

1 Dirac Electron

The size of the electron has been a long debated question. In classical physics the idea that the electron radius r_0 is purely electromagnetic leads to the calculation

$$r_0 = \frac{e^2}{mc^2} = 2.82 \times 10^{-13} \tag{1}$$

centimeters, while the electron's Compton radius

$$r_c = \frac{e^2}{\alpha mc^2} = \frac{e_*^2}{mc^2} = 3.86 \times 10^{-11} \tag{2}$$

is larger by the factor $1/\alpha\,(\approx 137)$, where $\alpha\,(= e^2/e_*^2)$ is the fine structure constant. The standard caveat at this point in the calculations is that, for any radius smaller than r_c (like r_0), classical considerations are irrelevant due to the possible appearance of electron-positron pairs. So the onset radius for electron-position pair production is an important parameter in the Dirac theory of the electron. What follows takes a detailed look at the structure of the second ratio in (2) and suggests that an onset radius derived from the coupling force the Dirac electron (De) exerts on the vacuum state produces a better estimate of that radius.

In the Planck vacuum (PV) theory [1] the product $e_*^2 = (-e_*)(-e_*)$ in (2) consists of two distinctively different charges. One of the bare charges belongs to the De (a massive point charge $(-e_*, m)$ that obeys the Dirac equation and that is coupled to the Dirac vacuum [2]), and the other to the separate Planck particles constituting the PV negative-energy state. In addition, it can be argued [3] that the force

$$\frac{e_*^2}{r^2} \tag{3}$$

is a polarization-distortion force that the free-space De exerts on the omnipresent PV state. Since this force exists between the electron charge and the individual Planck-particle charges within the PV, a potential

$$V(r) = -\int_{r_1}^{r} \frac{e_*^2}{r^2} dr = \left(\frac{1}{r} - \frac{1}{r_1}\right) e_*^2 \tag{4}$$

can be defined for the De-PV system, except for the difficulty in determining the integration constant r_1.

The massive point charge $(-e_*, m)$ has two parts, its charge $(-e_*)$ and its mass m. Thus, in addition to the polarization force (3), the De distorts the PV due to a gravitational-like attraction between its mass and the individual masses of the Planck particles in the PV. This curvature force is given by [3]

$$-\frac{mc^2}{r} = -\frac{mc^2 G}{r\,G} = -\frac{mm_* G}{r_* r} \tag{5}$$

where m_* and r_* are the mass and Compton radius of the individual Planck particles and G is Newton's gravitational constant. ($G = e_*^2/m_*^2$ and $e_*^2 = r_* m_* c^2$ are used in deriving the final ratio in (5).) This force is the force of attraction the massive point charge at $\langle r\rangle \approx 0$ exerts on the negative-energy Planck particle at a radius r from that charge. Now the total De distortion force becomes

$$\frac{e_*^2}{r^2} - \frac{mc^2}{r} \tag{6}$$

and, as seen in the next section, the r_1-problem of the previous paragraph disappears. Part of the response to the De force (6) acting on the PV is hidden in the Dirac equation as the zitterbewegung.

[The average $\langle r\rangle \approx 0$ signifies a small, but unknown, radius encircling the massive point charge $(-e_*, m)$ and in which the electron mass is created (see the Appendix). This average is more properly expressed as $\sqrt{\langle r^2\rangle} \ll r_c$.]

2 Dirac Equation

The force difference in (6) vanishes at the De's Compton radius

$$r_c = \frac{e_*^2}{mc^2} \tag{7}$$

which is that radius where the polarization and curvature forces have the same magnitude. This is a central parameter in the theory of the electron-positron system, for the free-

particle Dirac equation can be expressed as (using $c\hbar = e_*^2$) [4, p. 74]

$$ie_*^2\left(\frac{\partial}{c\partial t} + \boldsymbol{\alpha}\cdot\nabla\right)\psi = mc^2\beta\psi \qquad \text{or}$$

$$ir_c\left(\frac{\partial}{c\partial t} + \boldsymbol{\alpha}\cdot\nabla\right)\psi = \beta\psi \tag{8}$$

where, in the rest frame of the De, the parameter r_c represents the radius of an imaginary sphere surrounding the massive point charge and on which the PV is undistorted (where (9) and (10) vanish).

Now the De-PV coupling force

$$F(r) = \frac{e_*^2}{r^2} - \frac{mc^2}{r} \tag{9}$$

leads, in place of (4), to the potential

$$V(r) = -\int_{r_c}^{r} F(r)dr = \left(\frac{1}{r} - \frac{1}{r_c}\right)e_*^2$$

$$-mc^2\ln\frac{r_c}{r} \qquad (r \leqslant r_c) \tag{10}$$

with no undetermined constants.

Recalling that any sufficiently strong positive potential acting on the vacuum state enables electron-positron pair production to take place in free space (see any relativistic discussion of the Klein Paradox, e.g. [4, p. 131]), it is reasonable to conclude that the point at which pairs may begin to show up as the De is approached is where $V(r) = 2mc^2$ since the positive energy in free space and negative energy of the PV begin to overlap at this potential. Then solving (10) for r yields the quadrature formulas

$$\frac{r_c}{r} - \ln\frac{r_c}{r} = 3 \qquad \text{or} \qquad \frac{\exp(r_c/r)}{r_c/r} = e^3 \tag{11}$$

either one of which produces $r \approx r_c/4.5$. This pair-production onset radius is significantly smaller than the standard estimate ($r \sim r_c$) because the curvature-force term in (9) compresses the PV state, countering the polarization force that expands that state and exposes its energies to free space. This important result implies that, for any $r > r_c/4.5$, there can be no exchange of free electrons with electrons from electron-positron pairs associated with the PV state.

3 QED Comparison

The standard estimate of the onset radius is based on virtual electron-positron transitions and the time-energy uncertainty relation [5, p. 323]

$$\Delta t\,\Delta E \sim \hbar \quad \longrightarrow \quad c\Delta t \sim \frac{c\hbar}{\Delta E} = \frac{e_*^2}{2mc^2} = \frac{r_c}{2} \tag{12}$$

where the original free electron jumps into the positron hole and the electron from the pair becomes the new free electron. As this process takes place at a high rate, the resulting cloud of "hide-and-seek" electrons is perceived as a spread-out point electron with a radius $r \sim r_c/2$. This radius is usually rounded off to $r \sim r_c$. It is interesting that arbitrarily replacing r_1 in (4) by r_c leads to the estimate $r = r_c/3$.

Whatever the true magnitude of the onset radius, it is worth noting the following quantum electrodynamic conclusions [5, pp. 402–403]: the interaction of the De with the quantum vacuum spreads out the point-like nature of the De and leads to a natural scale r_c for the model; the De in some respects behaves as though it increases in size from a point particle to a particle with a radius of about one r_c; it is improbable that the electron has "structure"; and the apparent spread of the De does not alter the fact that the electron in QED is still regarded as a pure point particle. In addition to these conclusions, high-energy scattering experiments probing small distances indicate that the electron, if not a point particle, is certainly not larger than about 10^{-15} cm ($r_c/39,000$).

Except for the magnitude of the onset and spread radii, the calculations in Sections 1 and 2 are mostly in agreement with the spirit of the QED conclusions of the previous paragraph. Also the earlier assumption at the end of Section 1, that $\langle r\rangle \approx 0$, is in line with the experimental result ($r_c/39,000$) at the end of the previous paragraph.

Since the onset radius is an important concept in the electron model, a definitive calculation of this radius is crucial to understanding the electron — indeed, contrary to the standard view, it is shown in the present paper that the Compton radius r_c and the pair-creation onset radius $r_c/4.5$ are two *distinctly different* parameters, the first referring to the vanishing-coupling-force sphere centered on the point electron (in its rest frame), and the second to the possible onset of electron-positron pairs. This separation of the Compton and onset radii leads to a believable zitterbewegung model.

4 Zitterbewegung

The zitterbewegung (a highly oscillatory, microscopic motion with velocity c) has been a long-time mathematical conundrum. Barut and Bracken [6, p. 2458] reexamine the Schrödinger calculations leading to the zitterbewegung and replace his "microscopic momentum" vector with a "relative momentum" vector in the rest frame of the particle. Of interest here are the two resulting commutator brackets ($\hbar = r_c mc$ and $c\hbar = e_*^2$ are used)

$$[Q_j, H_r] = ir_c cP_j \qquad \text{and}$$

$$[P_j, H_r] = -4i\frac{mce_*^2}{r_c^2}\,Q_j \tag{13}$$

from the theory, where ($j = 1, 2, 3$) and $H_r = mc^2\beta$ is the Dirac Hamiltonian in the rest frame.

Applying the Heisenberg-picture time derivative

$$\dot{A} = \frac{i}{\hbar}[H_r, A] \tag{14}$$

to the commutators in (13) leads to the "relative momentum"

$$P_j = m\dot{Q}_j \qquad \text{and} \qquad \dot{P}_j = -4\left(\frac{e_*^2}{r_c^3}\right)Q_j \tag{15}$$

which describes the dynamics of a harmonic oscillator with angular frequency

$$\omega = \left(\frac{4\cdot e_*^2}{mr_c^3}\right)^{1/2} = \left(\frac{4\cdot r_c mc^2}{mr_c^3}\right)^{1/2} = \frac{2c}{r_c}\,. \tag{16}$$

Since the Compton relation derives from the equality of the polarization- and curvature-force magnitudes on the r_c-sphere surrounding the massive point charge, the oscillator dynamics must be due to a reaction of the PV to the De perturbing force $e_*^2/r^2 - mc^2/r$, *not* to a direct dynamical involvement of the massive point charge itself. This latter conclusion is supported by the fact that the eigenvalues of the $\dot{Q}_j$ operator are $\pm c$, outlawing the involvement of a massive particle whose velocity must be less that c.

The "spring constant", $4(e_*^2/r_c^3)$, in (15) is easily shown to be related to the r_c-sphere, for $r = r_c + \Delta r$ in (9) leads to

$$F(r_c + \Delta r) = \frac{e_*^2}{(r_c+\Delta r)^2} - \frac{mc^2}{r_c + \Delta r}$$

$$= -\frac{(e_*^2/r_c^3)\Delta r}{(1+\Delta r/r_c)^2} \approx -\left(\frac{e_*^2}{r_c^3}\right)\Delta r \tag{17}$$

where $F(r_c) = 0$, and $\Delta r \ll r_c$ in the final ratio.

The Schrödinger "microscopic coordinate"

$$\boldsymbol{\xi} = \left[\boldsymbol{\alpha}(0) - \frac{mc^2}{H}\frac{\widehat{\mathbf{p}}}{mc}\right]\cdot\frac{ir_c}{2}\frac{mc^2}{H}\exp\left[-i\frac{2c}{r_c}\frac{H}{mc^2}t\right] \tag{18}$$

is retained in the Barut-Bracken analysis [6, eqn. 19]. The first part of this operator equation corresponds to the macroscopic motion of the massive point charge and the second part to the high-frequency zitterbewegung superimposed on the macroscopic motion. In the rest frame of the massive charge (18) reduces to [6, eqn. 34]

$$Q_j(t) = [\boldsymbol{\xi}_r(t)]_j = \alpha_j(0)\cdot\frac{ir_c}{2}\beta\exp\left[-i\frac{2c}{r_c}\beta t\right] \neq 0 \tag{19}$$

the nonvanishing of which emphasizes again that the zitterbewegung is not fundamentally associated with the motion of the particle, as the particle leading to (19) is at rest. (The rest frame operators $H_r = mc^2\beta$ and $H_r^{-1} = \beta/mc^2$ are used in (19)).

5 Comments and Summary

The preceding calculations have separated the Compton radius (r_c) from the onset radius ($r_c/4.5$), with the result that the Compton radius is no longer associated with electron-positron pair production, being outside the onset radius. Thus the zitterbewegung is not related to the pair-production characteristic of an over-stressed ($V(r) \geqslant 2mc^2$) PV state. Instead, the zitterbewegung is seen to be the consequence of a PV-resonance phenomenon (with the resonant frequency $2c/r_c$) associated with the r_c-sphere. Also, most of the confusion surrounding the zitterbewegung is the result of attempting to attribute the phenomenon directly to the dynamics of the electron particle rather than the dynamics of the vacuum state. Finally, the zitterbewegung can now be seen, not as a mathematical curiosity, but as an integral part of the Dirac electron theory.

The following picture of the Dirac electron emerges: centered at the origin of the rest frame is the massive point charge with an effective volumetric radius $\langle r\rangle \approx 0$; surrounding this charge is a hypothetical sphere of radius $r_c/4.5$ within which the positive energy of the free electron and the negative energy of the PV overlap, allowing electron-positron pairs to be excited; surrounding this combination is a spherical annulus of radius $r_c/4.5 < r \leqslant r_c$, where pair production does not occur; and beyond the r_c-sphere ($r \geqslant r_c$) is a region of diminishing PV stress, a compression that decreases with increasing r according to the force difference (9).

Appendix: Electron Mass

The massless point charge is denoted by $(-e_*)$ and the massive point charge by $(-e_*, m)$, where m is the electron mass. In the PV theory this mass is an acquired property of the electron, resulting from the point charge being driven by the random electromagnetic zero-point background field [7, 8]. Furthermore, the energy absorbed by the charge from the field is re-radiated back into free space in a detailed-balance manner, leaving the isotropy and spectral density of the zero-point background unchanged [9].

The derived mass is

$$m = \frac{4r_c e_*^2}{9r_*^2c^2}\frac{\left\langle (dr'/dt)^2\right\rangle}{c^2} \tag{A1}$$

where e_*r' is the dipole moment of the point charge $(-e_*)$ about $r' = 0$ as it is being driven by the zero point field. The relative root-mean-square velocity of the charge within $\langle r\rangle \approx 0$ is [7]

$$\left(\frac{\left\langle (dr'/dt)^2\right\rangle}{c^2}\right)^{1/2} = \frac{3}{2}\frac{r_*}{r_c} \sim 10^{-22} \tag{A2}$$

which is vanishingly small because of the large density ($\sim 1/r_*^3$) of Planck particles in the PV contributing simultaneously to the zero-point background field; endowing the corresponding field spectrum with frequencies as high as $\sim c/r_*$,

where r_* is the Planck length. It is predominately the high frequencies in the spectrum that define the mass and prevent the r-m-s velocity from significantly increasing in magnitude [10].

The squared charge e_*^2 in (A1) comes from squaring the time derivative of the dipole moment $e_* r'$. Thus (A1) implies that *the center-of-mass and the center-of-charge are the same*. The question of centers often comes up in the discussion of the zitterbewegung [11, pp. 62–64] and is a reflection of the fact that the zitterbewegung is being explained in terms of the massive-charge motion rather than the $2c/r_c$ resonance associated with the r_c-sphere and the vacuum state.

Submitted on April 18, 2013 / Accepted on April 25, 2013

References

1. Daywitt W.C. The Planck vacuum. *Progress in Physics*, 2009, v. 1, 20–26.
2. Dirac P.A.M. A Theory of Electrons and Protons. *Proceedings of the Royal Society of London A*, 1930, v. 126, 360–365.
3. Daywitt W.C. The Dirac Electron in the Planck Vacuum Theory. *Progress in Physics*, 2010, v. 4, 69–71.
4. Gingrich D.M. Practical Quantum Electrodynamics. CRC, The Taylor & Francis Group, Boca Raton, London, New York, 2006.
5. Milonni P.W. The Quantum Vacuum – an introduction to quantum electrodynamics. Academic Press, New York, 1994.
6. Barut A.O. and Bracken A.J. *Zitterbewegung* and the internal geometry of the electron. *Physical Review D*, 1981, v. 23, no. 10, 2454–2463.
7. Daywitt W.C. The Source of the Quantum Vacuum. *Progress in Physics*, 2009, v. 1, 27–32.
8. Puthoff H.E. Gravity as a zero-point-fluctuation force. *Physical Review A*, 1989, v. 39, no. 5, 2333–2342.
9. Boyer T.H. Random electrodynamics: the theory of classical electrodynamics with classical electrodynamic zero-point radiation. *Physical Review D*, 1975, v. 11, no. 4, 790–808.
10. Daywitt W.C. Neutron Decay and its Relation to Nuclear Stability. To be published in *Galilean Electrodynamics* (pre-published in www.planckvacuum.com). See the appendix of this reference.
11. Grandy Jr. W.T. Relativistic Quantum Mechanics of Leptons and Fields. Kluwer Academic Publishers, Dordrecht—London, 1991.

Geometrical Derivation of the Lepton PMNS Matrix Values

Franklin Potter

Sciencegems.com, 8642 Marvale Drive, Huntington Beach, CA, USA. E-mail: frank11hb@yahoo.com

The linear superposition of generators of the 3 discrete binary rotational subgroups [332], [432], [532] of the Standard Model determine the PMNS matrix elements. The 6 leptons are 3-D entities representing these 3 groups, one group for each lepton family.

1 Introduction

Numerous attempts to derive the neutrino PMNS matrix from various discrete group horizontal symmetries have led to partial success. Herein I determine the true source of the PMNS matrix elements by using the linear superposition of the generators for 3 discrete binary rotational subgroups of the Standard Model (SM) electroweak gauge group $SU(2)_L$ x $U(1)_Y$.

In a series of articles [1–4] I have proposed 3 discrete binary rotational subgroups of the SM gauge group for 3 lepton families in R^3 and the related 4 discrete binary rotational subgroups in R^4 for 4 quark families, one binary group for each family. The generators for these 7 binary groups are quaternions operating in R^3, in R^4, and in C^2. I use these binary group quaternion generators to calculate the matrix elements for the PMNS mixing matrix for the leptons.

In another article under preparation I use the same approach, with an important modification, to calculate the standard CKM mixing matrix for the quarks as well as a proposed CKM4 mixing matrix for four quark families.

The SM local gauge group $SU(2)_L$ x $U(1)_Y$ x $SU(3)_C$ defines an electroweak(EW) interaction part and a color interaction part. The EW isospin states define the flavor of the fundamental lepton and quark states. However, experiments have determined that these left-handed flavor states are linear superpositions of mass eigenstates.

For the 3 lepton families, one has the neutrino flavor states ν_e, ν_μ, ν_τ and the mass states ν_1, ν_2, ν_3 related by the PMNS matrix U_{ij}

$$\begin{bmatrix}\nu_e\\ \nu_\mu\\ \nu_\tau\end{bmatrix} = \begin{bmatrix}U_{e1} & U_{e2} & U_{e3}\\ U_{\mu1} & U_{\mu2} & U_{\mu3}\\ U_{\tau1} & U_{\tau2} & U_{\tau3}\end{bmatrix}\begin{bmatrix}\nu_1\\ \nu_2\\ \nu_3\end{bmatrix}$$

From experiments [5], the PMNS angles have been estimated to be

$$\theta_{12} = 32.6^\circ - 34.8^\circ, \quad \theta_{13} = 8.5^\circ - 9.4^\circ,$$

$$\theta_{23} = 37.2^\circ - 39.8^\circ, \quad \delta = (0.77 - 1.36)\pi.$$

Consequently, for the normal hierarchy of neutrino masses, one has the empirically determined PMNS matrix

$$\begin{bmatrix} 0.822 & 0.547 & -0.150 + 0.038i\\ -0.356 + 0.0198i & 0.704 + 0.0131i & 0.614\\ 0.442 + 0.0248i & -0.452 + 0.0166i & 0.774\end{bmatrix}$$

which can be compared to my resultant derived PMNS matrix in the standard parametrization

$$\begin{bmatrix} 0.817 & 0.557 & -0.149e^{-i\delta}\\ -0.413 - 0.084e^{i\delta} & 0.605 - 0.057e^{i\delta} & -0.673\\ -0.383 + 0.090e^{i\delta} & 0.562 + 0.061e^{i\delta} & 0.725\end{bmatrix}$$

In the SM the EW isospin symmetry group that defines the lepton and quark flavor states is assumed to be the Lie group SU(2) with its two flavor eigenstates per family. In this context there is no fundamental reason for Nature to have more than one fermion family, and certainly no reason for having 3 lepton families and at least 3 quark families. As far as I know, this normal interpretation of the SM provides no answer that dictates the actual number of families, although the upper limit of 3 lepton families with low mass neutrinos is well established via Z^0 decays and via analysis of the CMB background. There are claims also that one cannot have more than 15 fundamental fermions (plus 15 antifermions) without violating certain cosmological constraints.

My geometrical approach makes a different choice, for I utilize discrete binary rotational subgroups of SU(2) instead, a different subgroup for each family. Each discrete binary group has two eigenstates and three group generators, just like SU(2). Whereas the three generators for the SU(2) Lie group are essentially the 2 x 2 Pauli matrices, the three generators for each of the 3 lepton discrete binary groups [332], [432], [532], (also labeled 2T, 2O, 2I) in R^3 and the 4 quark discrete groups [333], [433], [343], [533], (also labeled 5-cell, 16-cell, 24-cell, 600-cell) in R^4 are not exactly the Pauli matrices.

I propose that this difference between the discrete subgroup generators and the Pauli matrices is the fundamental source of the lepton and the quark mixing matrices, and the calculated results verify this conjecture. In other words, one requires the mixing of the different family discrete groups in order to have a complete set of three generators equivalent to the three SU(2) generators, separately for the leptons and for the quarks. The mixing matrices, PMNS and CKM4, express this linear superposition of the discrete group generators.

2 The PMNS calculation

In order to calculate the PMNS values one can use either unit quaternions or unitary 2x2 complex matrices. The unit quaternion generators are equivalent to the SU(2) generators.

The unit quaternion q = a + b**i** + c**j** + d**k**, where the coefficients a, b, c, d are real numbers for the one real and three imaginary axes. The unit quaternion spans the space R^4 while the imaginary prime part spans the subspace R^3. With $i^2 = j^2 = k^2 = -1$, the quaternion can be expressed as an SU(2) matrix

$$\begin{bmatrix} a+bi & c+di \\ -c+di & a-bi \end{bmatrix}$$

Both the quaternions and the SU(2) matrices operate in the unitary plane C^2 with its two orthogonal complex axes, so the quaternion can be written also as q = u + v**j**, with u = a + b**i** and v = c + d**i**. The three Pauli matrices σ_x, σ_y, σ_z, are the simple quaternions k, j, and i, respectively.

For the three lepton families, each family representing its own binary rotational group, [332], [432], and [532], two of the three generators R_i, i = 1, 2, 3, in each group are equivalent to two of the three Pauli matrices. Therefore, only the remaining generator for each lepton family contributes to the mixing that produces the PMNS matrix. That is, in the notation of H.M.S. Coxeter [6], R_1 = j, R_3 = i, and

$$R_2 = -i\cos\frac{\pi}{q} - j\cos\frac{\pi}{p} + k\sin\frac{\pi}{h} \qquad (1)$$

for the three binary groups [p q r] and the h values 4, 6, and 10, respectively.

Defining the golden ratio $\phi = (\sqrt{5}+1)/2$, the appropriate generators R_2 are listed in the table. The sum of all three R_2 generators should be k, so one has three equations for three unknowns, thereby determining the listed multiplicative factor for each R_2 generator's contribution to k after overall normalization.

Table 1: Lepton Family Discrete Group Assignments

Family	Group	R_2	Factor	Angle°
ν_e, e	[332]	$-\frac{1}{2}i - \frac{1}{2}j + \frac{1}{\sqrt{2}}k$	-0.2645	105.337
ν_μ, μ	[432]	$-\frac{1}{2}i - \frac{1}{\sqrt{2}}j + \frac{1}{2}k$	0.8012	36.755
ν_τ, τ	[532]	$-\frac{1}{2}i - \frac{\phi}{2}j + \frac{\phi^{-1}}{2}k$	-0.5367	122.459

The resulting angles in the table are determined by the arccosines of the factors, but they are twice the rotation angles required in R^3, a property of quaternion rotations. Using one-half these angles produces

$$\theta_1 = 52.67^\circ, \quad \theta_2 = 18.38^\circ, \quad \theta_3 = 61.23^\circ, \qquad (2)$$

resulting in

$$\theta_{12} = 34.29^\circ, \quad \theta_{13} = -8.56^\circ, \quad \theta_{23} = -42.85^\circ. \qquad (3)$$

Note that $|\theta_{12} - \theta_{13}| = |\theta_{23}|$ because of normalization.

Products of the sines and cosines of these angles in the standard parameterization are the PMNS entries, producing matrix values which compare favorably with the empirical estimates, as shown earlier. One has $\sin^2\theta_{12} = 0.3176$ and $\sin^2\theta_{13} = 0.0221$, both within 1σ of the empirically determined values from the neutrino experiments, according to the Particle Data Group in 2012. However, $\sin^2\theta_{23} = 0.4625$ is outside the PDG 1σ range but agrees with the recent T2K [7] estimate $\sin^2 2\theta_{23} = 1.0$, making $|\theta_{23}| = 45^\circ$ with $\delta \simeq 0$.

3 Conclusions

This fit of the PMNS mixing matrix derived from the three separate R_2 generators indicates that the lepton families faithfully represent the discrete binary rotational groups [332], [432], and [532] in R^3 that were introduced first in my geometrical approach back in 1986 and expanded in detail over the past two decades. In particular, the 6 lepton states are linear superpositions of the two degenerate basis states in each of the 3 groups. My approach within the realm of the Standard Model local gauge group makes the ultimate *unique* connection to the discrete group Weyl E_8 x Weyl E_8 in 10-D spacetime and to the Golay-24 code in information theory [1].

One can conclude that leptons are 3-dimensional objects, geometrically different from the quarks which require a 4-dimensional space for their existence. Their mass ratios derive from a mathematical syzygy relation to the j-invariant of elliptic modular functions associated with these specific binary groups. In addition, one can predict that no more lepton families exist because the appropriate binary rotational symmetry groups in 3-D space have been exhausted. However, sterile neutrinos remain viable [1, 4].

Acknowledgements

The author thanks Sciencegems.com for encouragement and financial support.

Submitted on April 22, 2013 / Accepted on April 29, 2013

References

1. Potter F. Our Mathematical Universe: I. How the Monster Group Dictates All of Physics. *Progress in Physics*, 2011, v. 4, 47–54.
2. Potter F. Discrete Rotational Subgroups of the Standard Model dictate Family Symmetries and Masses. DISCRETE'08 Conference, 2008. Online: www.sciencegems.com/DISCRETE08.PDF
3. Potter F. Unification of Interactions in Discrete Spacetime. *Progress in Physics*, 2006, v. 1, 3–9.
4. Potter F. Geometrical Basis for the Standard Model. *International Journal of Theoretical Physics*, 1994, v. 33, 279–305. Online: www.sciencegems.com/gbsm.html
5. Fogli G.L. et al. Global analysis of neutrino masses, mixings and phases. arXiv: 1205.5254v3.
6. Coxeter H.S.M. Regular Complex Polytopes. Cambridge University Press, Cambridge, 1974.
7. Abe K. et al. Evidence of Electron Neutrino Appearance in a Muon Neutrino Beam. arXiv: 1304.0841v1.

Higgs-Like Particle due to Revised Quantum Electrodynamics

Bo Lehnert

Alfvén Laboratory, Royal Institute of Technology, SE-10044 Stockholm, Sweden. E-mail: Bo.Lehnert@ee.kth.se

A Higgs-like particle having zero net electric charge, zero spin, and a nonzero rest mass can be deduced from an earlier elaborated revised quantum electrodynamical theory which is based on linear symmetry breaking through a nonzero electric field divergence in the vacuum state. This special particle is obtained from a composite longitudinal solution based on a zero magnetic field strength and on a nonzero divergence but a vanishing curl of the electric field strength. The present theory further differs from that of the nonlinear spontaneously broken symmetry by Higgs, in which elementary particles obtain their masses through an interaction with the Higgs field. An experimental proof of the basic features of a Higgs-like particle thus supports the present theory, but does not for certain confirm the process which would generate massive particles through a Higgs field.

1 Introduction

As stated in a review by Quigg [1] among others, the Higgs boson is a particle of zero electric charge and nonzero rest mass. The magnitude of the mass is, however, so far not predicted by theory. Several authors and recently Garisto and Argawal [2] have further pointed out that this particle is a spin-zero boson.

In this investigation will be shown that a particle with such basic properties can be deduced from an earlier elaborated revised quantum electrodynamical theory [3], and the consequences of this will be further discussed here.

2 Steady Axisymmetric States of Revised Quantum Electrodynamics

For the field equations of the revised theory to be used in this context, reference is made to earlier detailed deductions [3]. The latter are based on a broken symmetry between the field strengths $\mathbf{E}$ and $\mathbf{B}$, through the introduction of a nonzero divergence $\operatorname{div}\mathbf{E}=\bar{\rho}/\varepsilon_0$ as being based on the quantum mechanical Zero Point Energy of the vacuum state. In a spherical frame (r,θ,φ) of reference in an axially symmetric steady state with $\partial/\partial\varphi=0$ and $\partial/\partial t=0$, this leads to a magnetic vector potential $\mathbf{A}=(0,0,A)$ and a space charge current density $\mathbf{j}=(0,0,C\bar{\rho})$ due to the source $\bar{\rho}$. Here $C=\pm c$ represents the two spin directions, with c standing for the velocity constant of light. Introducing the normalized radius $\rho=r/r_0$ with r_0 as a characteristic length, and the separable generating function

$$F(r,\theta)=CA-\phi=G_0G(\rho,\theta)=G_0R(\rho)\cdot T(\theta) \qquad (1)$$

where ϕ is the electrostatic potential, this yields

$$CA = -(\sin\theta)^2\,DF \qquad (2)$$

$$\phi = -\left[1+(\sin\theta)^2\,D\right]F \qquad (3)$$

$$\bar{\rho} = -\frac{\varepsilon_0}{r_0^2\rho^2}D\left[1+(\sin\theta)^2\,D\right]F \qquad (4)$$

with the operator

$$\begin{aligned} D &= D_\rho + D_\theta \\ D_\rho &= -\frac{\partial}{\partial\rho}\left(\rho^2\frac{\partial}{\partial\rho}\right) \\ D_\theta &= -\frac{\partial^2}{\partial\theta^2}-\frac{\cos\theta}{\sin\theta}\frac{\partial}{\partial\theta}. \end{aligned} \qquad (5)$$

The field strengths then become

$$\mathbf{B}=\operatorname{curl}\mathbf{A}=\operatorname{curl}\left[0,0,-\frac{1}{C}(\sin\theta)^2\,DF\right] \qquad (6)$$

$$\mathbf{E}=-\nabla\phi=\nabla\left\{\left[1+(\sin\theta)^2\,D\right]F\right\} \qquad (7)$$

for an elementary mode generated by a given function F determined by the radial and polar parts $R(\rho)$ and $T(\theta)$. Here $\operatorname{curl}\mathbf{E}=0$.

As a first step we consider the convergence properties of R and the symmetry properties of T with respect to the equatorial plane $\theta=\pi/2$. There are four alternatives of which there is one with a divergent R at the origin $\rho=0$ and with a T of top-bottom symmetry, thereby leading to a net integrated electric charge q_0 and magnetic moment M_0. The other three alternatives all lead to vanishing q_0 and M_0 [3], and we can here choose any of these. Then the local electric field and its divergence are still nonzero, whereas the net integrated electric charge vanishes.

As a second step two elementary modes $(^+)$ and $(^-)$ are now considered for which $C=\pm c$ and there is the same function F. For these modes the corresponding field strengths are related by

$$\mathbf{B}^+=-\mathbf{B}^- \qquad \mathbf{E}^+=\mathbf{E}^- \qquad (8)$$

according to Equations (6) and (7). Since the field equations are linear, the sum of the two solutions $(^+)$ and $(^-)$ also becomes a solution of the field equations, thereby resulting in the field strengths

$$\mathbf{B}_H=\mathbf{B}^++\mathbf{B}^-=0 \qquad (9)$$

$$\mathbf{E}_H = \mathbf{E}^+ + \mathbf{E}^- = \nabla\Big\{\Big[1 + (\sin\theta)^2\, D\Big]\, 2F\Big\}. \qquad (10)$$

These strengths then stand for a "composite" mode having zero integrated charge q_o, zero magnetic moment M_0, and zero spin s_0, but a nonzero rest mass

$$m_0 = \frac{4\pi\varepsilon_0}{c^2} r_0 G_0^2 J_m = \left(\frac{1}{2}\varepsilon_0 E_{eq}^2\right)\left(\frac{4}{3}\pi r_0^3\right). \qquad (11)$$

Here $E_{eq}^2 = 6(G_0/r_0)^2 J_m$, the dimensionless integral is

$$J_m = \int_0^{\infty}\int_0^{\pi} I_m \, d\rho \, d\theta \qquad I_m = fg \qquad (12)$$

and

$$f(\rho,\theta) = -(\sin\theta)\, D\Big[1 + (\sin\theta)^2\, D\Big] G \qquad (13)$$

$$g(\rho,\theta) = -\Big[1 + 2(\sin\theta)^2\, D\Big] G \qquad (14)$$

when a convergent radial part R is now being chosen [3].

It has first to be observed that this composite mode can be related to an option of the Higgs boson which is not truly a fundamental particle but is built out of as yet unobserved constituents, as also stated by Quigg [1]. Moreover, the vanishing magnetic field $\mathbf{B}_H$ of Equation (9) is in a way related to the longitudinal "S-wave" of the earlier theory [3], as well as to the longitudinal state of a massive boson mentioned by Higgs [4]. Finally, the magnitude of the nonzero mass is so far not predicted by theory [1]. From Equation (11) it should be due to the energy density of an equivalent electric field E_{eq}. The absence of a magnetic field may also make the particle highly unstable.

3 Discussion

An experimental proof of an existing Higgs-like particle with zero net electric charge, zero spin, and nonzero rest mass could thus be taken as support of the present revised quantum electrodynamical theory [3]. The latter is characterized by intrinsic linear symmetry breaking, leading in general to nonzero rest masses of elementary particles.

Such a proof does on the other hand not for certain become a full experimental confirmation also for the same particle to provide all other elementary particles with mass through the completely different spontaneous nonlinear symmetry breaking interaction between the Higgs field and massless particle concepts of the Standard Model [1, 4].

Possible the present approach [3] and that of Higgs [1, 4] could have a point in common. This is trough the Zero Point Energy field being present all over space on one hand [3], and a generally existing Higgs field in space on the other [1, 4].

Submitted on April 28, 2013 / Accepted on April 29, 2013

References

1. Quigg C. The coming revolution in particle physics. *Scientific American*, February 2008, 38–45.
2. Garisto R. and Agarwal A. The import of the Higgs boson. *Scientific American*, September 2012, page 16.
3. Lehnert B. Revised quantum electrodynamics. Nova Science Publishers, New York 2013, Chapter 6.
4. Higgs P.W. Spontaneous symmetry breakdown without massless bosons. *Physical Review*, 1966, v. 145, 1156–1163.

Key to the Mystery of Dark Energy: Corrected Relationship between Luminosity Distance and Redshift

T. X. Zhang

Department of Physics, Alabama A & M University, Normal, Alabama 35762. E-mail: tianxi.zhang@aamu.edu

A new possible explanation to the luminosity distance ($D_{\rm L}$) and redshift (Z) measurements of type Ia supernovae (SNeIa) is developed. Instead of modifying the theory of general relativity or the Friedmann equation of cosmology with an extra scalar field or unknown energy component (e.g., dark energy), we re-examine the relationship between the luminosity distance and the cosmological redshift ($D_{\rm L} - Z$). It is found that the $D_{\rm L} - Z$ relation previously applied to connect the cosmological model with the measured SNeIa data is only valid for nearby objects with $Z \ll 1$. The luminosity distances of all distant SNela with $Z \gtrsim 1$ had been underestimated. The newly derived $D_{\rm L} - Z$ relation has an extra factor $\sqrt{1+Z}$, with which the cosmological model exactly explains all the SNeIa measurements without dark energy. This result indicates that our universe has not accelerated and does not need dark energy at all.

1 Introduction

There are five possible ways to explain the luminosity distance ($D_{\rm L}$) and redshift (Z) measurements of type Ia supernovae (SNeIa) according to the general relativity (GR), which derives the Friedmann equation (FE) with the Friedmann-Lemaitre-Robertson-Walker (FLRW) metric of the 4D spacetime (Figure 1).

The most simple and direct way is the famous Lambda Cold Dark Matter (ΛCDM) model, currently accepted as the standard one, which introduces a cosmological constant Λ to the field equation of GR (Eq. 1), referred as a candidate of dark energy [1-2],

$$G_{\mu\nu} + \Lambda g_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu}, \qquad (1)$$

where $G_{\mu\nu}$ is the Einsteinian curvature tensor of spacetime, $T_{\mu\nu}$ is the energy-momentum tensor of matter, c is the light speed in free space, and G is the gravitational constant. The cosmological constant Λ was first introduced actually by Albert Einstein himself into his field equation, Eq. (1), in order to have a static universe about a century ago, and then discarded after the universe was found to be expanding [3].

The second way that has also been comprehensively studied is the scalar-tensor (S-T) theory, which introduces a scalar field Φ, usually time-dependent, to the action of spacetime (S_G) [4-5]. This category includes also the four-dimensional $f(R)$, galileon, and five-dimensional Kaluza-Klein theories with scalar fields [6-12]. The third way is the scalar perturbation (SP) theory, which inputs perturbation scalars Ψ and Φ, usually time-independent, into the FLRW metric rather than into the action S_G [13-15]. The S-T and SP theories may be equivalent because both attempt to modify the curvature of spacetime. The cosmological constant Λ can also be added to S_G for a less curvature of spactime or to the action of matter S_M for an extra energy component. The fourth possible way is according the black hole universe (BHU) model, recently developed by the author [16-18], in which the expansion and acceleration of the universe are driven by the external energy.

The procedures that the above four models commonly follow in the explanation of the SNeIa measurements include the following four steps: (1) Modifying the FE with an appropriate input of Λ, scalar field, perturbation, or external energy; (2) Determining the expansion rates (Hubble parameter) of the universe according to their modified FEs; (3) Submitting their expansion rates into the $D_{\rm L} - Z$ relation; (4) Comparing the obtained redshift dependence of their luminosity distances with the SNeIa measurements. Fitting the models to the data determines the amount of the input such as $\Omega_\Lambda \sim 0.73$ for the ΛCDM model [1-2] and $\ddot{M}(t) \sim 10^{17}$ kg/s^2 for the BHU model [18].

In this paper, a new and most probable explanation for the SNela measurements is developed without attempting to modify the theory of gravitation or the model of cosmology by inserting one or more fields or constants into GR or FE. Instead, we will re-examine the $D_{\rm L} - Z$ relationship that connects the cosmological model with the SNela data. We will derive a new $D_{\rm L} - Z$ relation and further compare this new relation with the SNela measurements to examine whether or not our universe needs the dark energy or has recently accelerated.

2 Mystery of Dark Energy

The greatest unsolved problem in the modern cosmology is the mystery of dark energy [19]. This currently most accepted hypothesis for the standard cosmological model to quantitatively explain the measurements of distant type Ia supernovae strongly relies on the $D_{\rm L} - Z$ relation that is used to bridge the measured SNeIa data and the theoretical model of cosmology.

However, the $D_{\rm L} - Z$ relation that was usually applied to analyze the measurements of distant type-Ia supernovae,

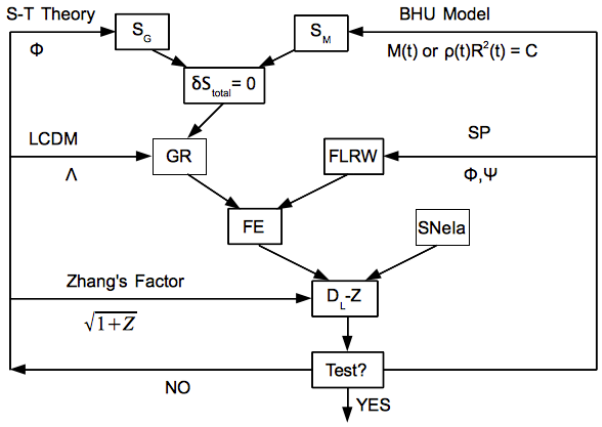

Fig. 1: Flow chat for five possible ways to explain the luminosity distance and redshift measurements of type Ia supernovae. They are: (1) GR with the cosmological constant Λ; (2) Gravitational theory with a scalar field Φ; (3) FLRW metric with perturbations Φ and Ψ; (4) Black hole universe model with increasing input of external energy $\dot{M} > 0$; and (5) Luminosity distance-redshift relation with a factor of $\sqrt{1+Z}$. This Study focuses on the fifth possible explanation.

$$D_{\rm L} \simeq c(1+Z)R(t_o)\int_{t_e}^{t_o}\frac{dt}{R(t)}, \tag{2}$$

is an approximate expression that is only valid for nearby objects with $Z \ll 1$ in a flat universe [20]. Here t_e is the time when the light is emitted, t_o is the time when the light is observed, $R(t)$ is the scale factor, which is defined from the FLRW metric [21-24],

$$ds^2 = -c^2dt^2 + R^2(t)\left[\frac{dr^2}{1-kr^2} + r^2\left(d\theta^2 + \sin^2\theta d\phi^2\right)\right], \tag{3}$$

and governed by the Friedmann equation [25],

$$H^2(t) \equiv \frac{\dot{R}^2(t)}{R^2(t)} = \frac{8\pi G\rho_M(t)}{3} - \frac{kc^2}{R^2(t)} + \frac{\Lambda}{3}, \tag{4}$$

according to the standard cosmological model, where $\rho_M(t)$ is the matter density, k is the curvature ($k = 0$ for a flat universe), Λ is the cosmological constant (or a candidate of dark energy), the coordinates $\{t, r, \theta, \phi\}$ are co-moving coordinates, and $H(t)$ is the Hubble parameter, which, at the present time, is called the Hubble constant and measured at $H_0 \sim 70$ km/s/Mpc [3, 26-27].

In the FLRW universe due to the time dependent scalar factor, light gets redshifted. According to the theory of GR, light travels on null geodesics (i.e., $ds^2 = 0$). Then along a radial light path, we have

$$\frac{cdt}{R(t)} = \frac{dr}{\sqrt{1-kr^2}}. \tag{5}$$

It follows from Eq. (5) that

$$\int_{t_e}^{t_o}\frac{cdt}{R(t)} = \int_{t_e+\delta t_e}^{t_o+\delta t_o}\frac{cdt}{R(t)} = \int_{r_1}^{0}\frac{dr}{\sqrt{1-kr^2}}. \tag{6}$$

Subtracting the first integral from the second and assuming $\delta t_e, \delta t_o << R(t)/\dot{R}(t)$, we get

$$\frac{\delta t_e}{R(t_e)} = \frac{\delta t_o}{R(t_o)}. \tag{7}$$

Since $\delta t_e = 1/\nu_e = \lambda_e/c$ and $\delta t_o = 1/\nu_o = \lambda_o/c$, the cosmological redshift Z can be determined according to the scale factor $R(t)$ as

$$1 + Z \equiv \frac{\lambda_o}{\lambda_e} = \frac{\nu_e}{\nu_o} = \frac{\delta t_o}{\delta t_e} = \frac{R(t_o)}{R(t_e)}. \tag{8}$$

Here λ and ν are the light wavelength and frequency, respectively. Light from a source object is redshifted because the time interval or scale factor is increased. The reason for an individual photon to be observed with smaller frequency (or energy) is due to that the time interval of observation is greater.

The scale factor is related to the energy and curvature via Eq. (4) and to the redshift via Eq. (8). In terms of Eqs. (4) and (8), the luminosity distance-redshift relation Eq. (2) can be reformed as

$$\begin{aligned} D_{\rm L} &\simeq c(1+Z)\int_0^Z\frac{dz'}{H(z')} \\ &= \frac{c}{H_0}(1+Z)\int_0^Z\frac{dz'}{\sqrt{\Omega_M(1+z')^3+\Omega_\Lambda}}, \end{aligned} \tag{9}$$

with $1 = \Omega_M + \Omega_\Lambda$. For an arbitrary k, Eq. (9) is generally represented as

$$\begin{aligned} D_{\rm L} \simeq\ & \frac{c}{H_0\sqrt{|\Omega_k|}}(1+Z)S\left(\sqrt{|\Omega_k|}\ \times\right. \\ & \left.\int_0^Z\frac{dz'}{\sqrt{\Omega_M(1+z')^3+\Omega_k(1+z')^2+\Omega_\Lambda}}\right), \end{aligned} \tag{10}$$

where

$$S(x) = \begin{cases} \sin(x), & \text{if } k < 0 \\ x, & \text{if } k = 0 \\ \sinh(x), & \text{if } k > 0 \end{cases} \tag{11}$$

and $1 = \Omega_M + \Omega_\Lambda + \Omega_k$.

Comparing the luminosity distance and redshift measurements of distant SNeIa with the luminosity distance-redshift relation determined in terms of Eqs. (2), (4), and (8) or Eq. (9) or Eq. (10) with $k = 0$, two supernova research groups [1-2], respectively, claimed that the universe has recently accelerated, so that the universe is dominated ($\Omega_\Lambda \sim 0.73$) by the dark energy.

However, re-examining the derivation of the luminosity distance-redshift relation, Eq. (2) so that Eqs. (9) and (10),

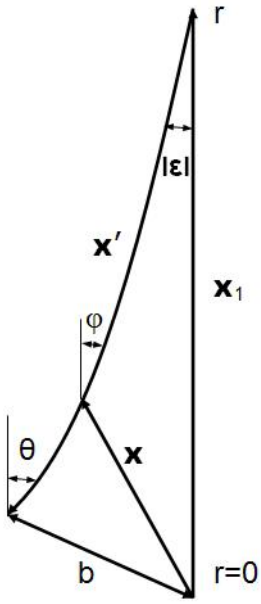

Fig. 2: Quantities used in the calculation of parallaxes and apparent luminosities [20]. The angles and the curvature of the light ray are greatly exaggerated.

we find that this relation is just an approximate relation only valid for nearby objects with $Z \ll 1$. Certainly, we cannot use it to correctly figure out the measurements of distant type Ia supernovae with $Z \gtrsim 1$. In the following, we will derive a new, more accurate and applicable also to distant objects, luminosity distance-redshift relation, which is perfectly consistent with all the measurements of type Ia supernovae without the input dark energy.

3 New $D_L - Z$ Relation

Now, the luminosity distance-redshift relation is derived by following the standard method as shown by [20] that calculates the parallaxes and apparent luminosities according to the path of light rays that leave from a source at $t = t_e$ and $r = r_e = R(t_e)r_1$ and pass to the observer at $t = t_o$ near $r = 0$ (see Figure 2). At the observation time $t = t_o$, the light source locates at $r = r_o = R(t_o)r_1$. Here, r_1 is the commoving distance defined by

$$r_1 = c\int_{t_e}^{t_o}\frac{dt}{R(t)}, \tag{12}$$

from the FLRW metric.

In the coordinate system x'^μ in which the light source is at the origin, the ray path is given by a position vector

$$\vec{x'} = \vec{n}\rho, \tag{13}$$

where $\vec{n}$ is a fixed unit vector and ρ is a variable positive parameter describing positions along the path. The coordinate system x'^μ can be transformed to another coordinate system x^μ in which the observer is at the origin (e.g., the center of the telescope) and the light source is at $\vec{x}_1$. In the observer coordinate system, the ray path can be represented by (Eq. 14.4.2 of [20])

$$\vec{x} = \vec{x'} + \vec{x}_1\left(\left(1-kx'^2\right)^{1/2} - \left\{1-\left(1-kx_1^2\right)^{1/2}\right\}\frac{(\vec{x}\cdot\vec{x}_1)}{x_1^2}\right). \tag{14}$$

For a flat universe ($k = 0$), the ray path in the coordinate system (Eq. 14) can be simplified as

$$\vec{x} = \vec{x'} + \vec{x}_1. \tag{15}$$

The parametric equation of the ray path, given by substituting Eq. (13) in Eq. (15), is then

$$\vec{x}(\rho) = \vec{n}\rho + \vec{x}_1. \tag{16}$$

The distance of light ray to the origin in the observer coordinate system will be

$$\begin{aligned} |\vec{x}| &= \sqrt{\left(\vec{x'} + \vec{x}_1\right)\cdot\left(\vec{x'} + \vec{x}_1\right)} \\ &= \sqrt{x_1^2 + \rho^2 - 2x_1\rho\cos\phi} \\ &\sim \sqrt{(x_1-\rho)^2 + x_1\rho\phi^2}, \end{aligned} \tag{17}$$

where we have considered the angle ϕ between $\vec{n}$ and $-\vec{x}_1$ is small and thus $\cos\phi \sim 1 - \phi^2/2$.

At the emission time t_e, we have

$$\rho|_{t=t_e} = 0, \tag{18}$$

$$|\vec{x}|_{t=t_e} = |\vec{x}_1|_{t=t_e} = r_e = r_1R(t_e), \tag{19}$$

$$\phi|_{t=t_e} = |\vec{\epsilon}|, \tag{20}$$

while at the observation time t_o, we have

$$\rho|_{t=t_0} = r_o = r_1R(t_o), \tag{21}$$

$$|\vec{x}|_{t=t_o} = b, \tag{22}$$

$$|\vec{x}_1|_{t=t_o} = r_o = r_1R(t_o), \tag{23}$$

$$\phi|_{t=t_o} = \theta = |\vec{\epsilon}|R(t_o)/R(t_e). \tag{24}$$

Substituting the quantity properties Eqs. (21)-(24) at $t = t_o$ into Eq. (17), we obtain the impact parameter as

$$b = R(t_o)r_1\theta = \frac{R^2(t_o)}{R(t_e)}r_1|\vec{\epsilon}|. \tag{25}$$

To calculate apparent luminosities, we consider a circular telescope mirror of radius b, placed with its center at the origin and its normal along the line of sight to the light source.

The fraction of all emitted photons that reach the mirror is the ratio of the solid angle to 4π,

$$\frac{\pi|\vec{\epsilon}|^2}{4\pi} = \frac{\pi b^2}{4\pi r_1^2}\frac{R^2(t_e)}{R^4(t_o)}. \quad (26)$$

Since light is red-shifted, the energy or frequency of each photon observed is reduced in comparison with the photon emitted by a factor of $R(t_e)/R(t_o)$. This energy or frequency reduction is equivalent to the increase of the time interval for observation relative to that for emission. If the effect of the redshift on the apparent luminosity is considered, then we should not consider the effect of the time interval increase on the apparent luminosity. This is also consistent with the electromagnetic wave theory of light, from which the energy emitted per unit time of emission is only one redshift factor greater than the energy observed per unit time of observation. Therefore, the total power P received by the mirror is the total power emitted by the source, its absolute luminosity L, times a factor $R(t_e)/R(t_o)$, and times the fraction (Eq. 26):

$$P = L\frac{\pi b^2}{4\pi r_1^2}\frac{R^3(t_e)}{R^5(t_o)}. \quad (27)$$

The apparent luminosity l is the power per unit mirror area

$$l = \frac{P}{\pi b^2} = \frac{L}{4\pi r_1^2}\frac{R^3(t_e)}{R^5(t_o)}. \quad (28)$$

Then the luminosity distance can be obtained

$$D_{\rm L} = \left(\frac{L}{4\pi l}\right)^{1/2} = r_1 R(t_o)\left[\frac{R(t_o)}{R(t_e)}\right]^{3/2}$$
$$= c(1+Z)^{3/2}R(t_o)\int_{t_e}^{t_o}\frac{dt}{R(t)}. \quad (29)$$

The luminosity distance Eq. (29) derived here is $\sqrt{1+Z}$ times that we conventionally used, Eq. (2). This factor leads to an explanation of type Ia supernova measurements without dark energy. Using Eqs. (4) and (8) for a flat universe ($k = 0$) without dark energy ($\Lambda = 0$), we can integrate Eq. (29) and obtain the luminosity distance-redshift relation as

$$D_{\rm L} = \frac{2c}{H_0}(1+Z)\left(\sqrt{1+Z}-1\right). \quad (30)$$

Eq. (30) dees not include any free parameter and reduces to the Hubble law at $Z \ll 1$.

The two significant corrections, which have been made in the above derivation of the luminosity distance in comparison with the derivation done in [20] are: 1) θ is not about equal to $|\vec{\epsilon}|$ for a distant light source but increased by a factor $R(t_o)/R(t_e)$, and 2) the light is red-shifted and the time interval increases are equivalent in physics $\nu_o/\nu_e = \delta t_e/\delta t_o = R(t_e)/R(t_o)$ and thus they reduce the apparent luminosity only by $R(t_e)/R(t_o)$ rather than its square. This is also supported by the electromagnetic wave theory of light.

The early derivation, including the simplified version as given in [28] and other cosmological books, the fraction of the light received in a telescope of aperture πb^2 on earth is $\pi b^2/[4\pi r_1^2 R^2(t_o)]$ and so the factor $1/d^2$ in the formula for the apparent luminosity l was replaced by $1/[r_1^2 R^2(t_o)]$. This replacement or modification for the apparent luminosity l was made according to the view of the emitter rather than from the view of the observer. From the view of the emitter (or a person standing on the source object), all light rays radially diverge from the source object isotropically and in straight lines. All the photons emitted at t_e reach the surface of the sphere drawn around the source object by radius $r_1 R(t_o)$. The angle of emission of a photon from the source object is equal to the angle of incidence of the photon to the mirror of telescope.

From the view of the observer, however, the source object is moving away in an increasing speed. The light rays travel in curved lines and anisotropically as shown in Figure 2. The angle of emission of a photon from the source object $|\vec{\epsilon}|$ is smaller than the angle of incidence of the photon to the mirror of telescope θ by a factor of $R(t_e)/R(t_o)$. That is, from the view of the observer, the factor $1/d^2$ in the formula for l must be replaced with $1/[r_1^2 R^4(t_o)/R^2(t_e)]$ as shown in Eq. (26). On the other hand, according the electromagnetic wave theory, the energy of radiation does not depend on the frequency. Only the increase of time interval would reduce the apparent luminosity. This may be examinable in experiments using a sound wave.

Figure 3 plots the luminosity distance-redshift relation (red line) along with the type Ia supernova measurements (blue dots. Credit: Union 2.1 compilation of 580 SNIA data from Supernova Cosmology Project). In this plot the Hubble constant is chosen to be $H_0 \sim 70$ km/s/Mpc. In the upper panel of Figure 3, the distance modulus, which is defined by $\mu = 5\log_{10} D_{\rm L} - 5$ with $D_{\rm L}$ in parsecs, is plotted as a function of redshift; while in the lower panel of Figure 3, the distance modulus difference between the measured SNeIa data and analytical results derived from Eq. (30). The chi-square statistic is obtained as

$$\chi^2 = \sum_{j=1}^{580}\frac{\left(\mu_j^{\rm obs} - \mu_j^{\rm the}\right)^2}{\sigma_j^2} \sim 589. \quad (31)$$

Then the reduced chi-square is given by $\chi^2_{\rm red} = 589/580 \sim 1.015$. It is seen that the derived luminosity distance-redshift relation is perfectly consistent with the measurements of type Ia supernovae. Therefore, with the new luminosity distance-redshift relation, the SNeIa measurements do not show the existence of dark energy.

The analysis and measurements for the structure and weak lensing of the CMB might not be accurate enough as were thought to provide an independent check or evidence on the

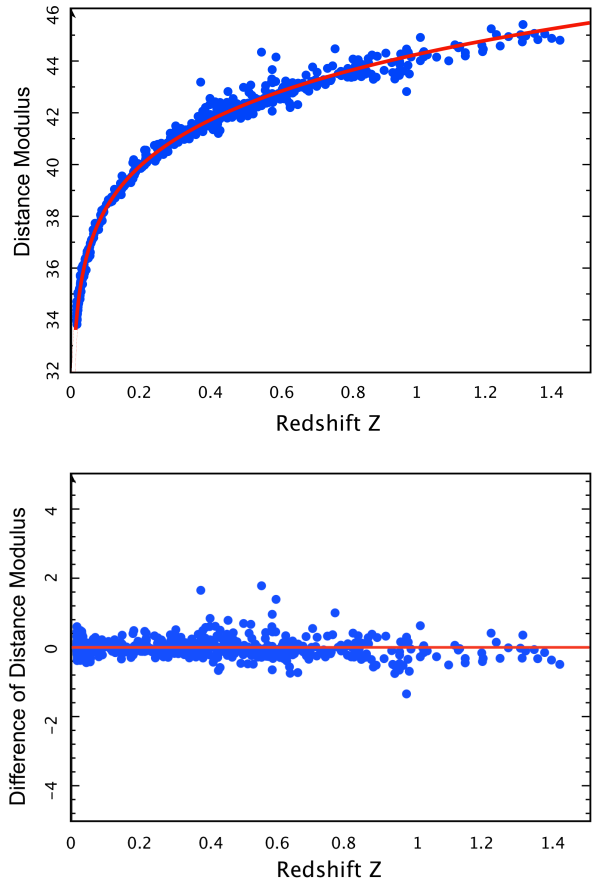

Fig. 3: Luminosity distance-redshift relation of type Ia supernovae. Blue dots are measurements credited by the Union2 compilation of 580 SNeIa data from Supernova Cosmology Project. Red lines are analytical results from this study. The upper panel plots the distance modulus as a function of redshift, while the lower panel plots the distance modulus difference between the measurement data and theoretical results.

existence of dark energy [29-30]. Recently, Sawangwit and Shank [31-32] looked at the CMB observations and find the errors in the data to be much larger than previously thought. The CMB power spectrum is very sensitive to the beam profiles. If their results are further confirmed to be correct, then it will also become less likely that dark energy dominates the universe.

4 Summary

The luminosity distance-redshift relation that we previously applied to connect the models with the SNeIa measurements is an approximate expression only valid for nearby objects. This is because that the traditional derivation of the $D_{\rm L}-Z$ relation has the following two defects: (1) the light emitting angle is about equal to the light incident angle, which is not true for light from a distant source object according to the view of the observer on the earth, and (2) the redshift of light and the increase of time interval doubly reduce the energy flux of the received light, which is physically incorrect because the redshift of light is caused by the increase of time interval. The electromagnetic wave theory of light also supports that the apparent luminosity is reduced only by one redshift factor due to the time interval increase. We have corrected these defects and derived a new relationship between luminosity distance and redshift with a factor of $\sqrt{1+Z}$. With this new $D_{\rm L}-Z$ relation, we have perfectly explained the SNela measurements according to the standard cosmological model without dark energy ($\Lambda=0$). Therefore, we can conclude that the universe has not accelerated and does not need the dark energy at all. The luminosity distance-redshift relation often used previously is only valid for nearby objects and thus the luminosity distances of all distant type Ia supernovae had been underestimated. This study provides us a possible solution to the mystery of dark energy.

Acknowledgements

This work was supported by the NASA EPSCoR (NNX-07AL52A) and the National Natural Science Foundation of China (G40890161).

Submitted on April 26, 2013 / Accepted on May 5, 2013

References

1. Riess A.G. et al. Observational evidence from supernovae for an accelerating universe and a cosmological constant. *Astronomical Journal*, 1998, v. 116, 1009–1038.
2. Perlmutter S. et al. Measurements of Omega and Lambda from 42 High-Redshift Supernovae. *Astrophysical Journal*, 1999, v. 517, 565–586.
3. Hubble E. A relation between distance and radial velocity among extragalactic nebulae. *Proceedings of the National Academy of Sciences of the United States of America*, 1929, v. 15, 168–173.
4. Brans C.H., Dicke R.H. Mach's principle and a relativistic theory of gravitation. *Physical Review* 1961, v. 124, 925–935.
5. Riazuelo A., Uzan J.P. Cosmological observations in scalar-tensor quintessence. *Physical Review D*, 2002, v. 66, 023525.
6. Buchdahl H.A. Non-linear Lagrangians and cosmological theory. *Monthly Notices of the Royal Astronomical Society*, 1970, v. 150, 1–8.
7. Sotiriou T.P., Faraoni V. f(R) theories of gravity. *Reviews of Modern Physics*, 2010, v. 82, 451–497.
8. Chow N., Khoury J. Galileon cosmology. *Physical Review D*, 2009, v. 80, 024037.
9. Freund P.G.O. Kaluza-Klein cosmologies. *Nuclear Physics B*, 1982, v. 209, 146–156.
10. Wesson P.S., Liu H. and Ahluwalia D.V. The cosmological constant problem and Kaluza-Klein theory. *International Journal of Modern Physics D*, 2001, v. 10, 905–912.
11. Zhang T.X. Electric redshift and quasars. *Astrophysical Journal Letters*, 2006, v. 636, L61–L64.
12. Zhang T.X. Gravitational field shielding and supernova explosions. *Astrophysical Journal Letters*, 2010a, v. 725, L117–L121.
13. Mukhanov V.F., Feldman H.A., Brandenberger R.H. Theory of cosmological perturbations. *Physics Reports*, 1992, v. 215, 203–333.

14. Bernardeau F., Colombi S., Gaztanaga E., Scoccimarro R. Large-scale structure of the Universe and cosmological perturbation theory. *Physics Reports*, 2002, v. 367, 1–248.
15. Wands D., Malik K.A., Lyth D.H., Liddle A.R. New approach to the evolution of cosmological perturbations on large scales. *Physical Review D*, 2000, v. 62, 043527.
16. Zhang T.X. A new cosmological model: black hole universe. *Bulletin of the American Astronomical Society*. 2007, v. 39, 1004–1004.
17. Zhang T.X. A new cosmological model: black hole universe. *Progress in Physics*, 2009, v. 2, 3-11.
18. Zhang T.X. Cosmic microwave background radiation and black hole universe. *Astrophysics and Space Science*, 2010b, v. 330, 157–165.
19. Peebles P.J. and Ratra B. The cosmological constant and dark energy. *Reviews of Modern Physics*, 2003, v. 75, 559–606.
20. Weinberg S. Gravitation and Cosmology, John Wiley & Sons, Inc., 1972.
21. Friedmann A. Über die Krümmung des Raumes. *Zeitschrift für Physik*, 1922, v. 10, 377–386.
22. Lemaitre G. Expansion of the universe, A homogeneous universe of constant mass and increasing radius accounting for the radial velocity of extra-galactic nebulae. *Monthly Notices of the Royal Astronomical Society*, 1931, v. 91, 483–490.
23. Robertson H.P. Kinematics and World-Structure. *Astrophysical Journal*, 1935, v. 82, 284–301.
24. Walker A.G. On Milne's theory of world-structure. *Proceedings of the London Mathematical Society*, 1937 v. 42, 90–127.
25. Friedmann A. Über die Möglichkeit einer Welt mit konstanter negativer Krümmung des Raumes. *Zeitschrift für Physik*, 1924, v. 21, 326–332.
26. Suyu S.H. et al. Dissecting the Gravitational lens B1608+656. II. Precision Measurements of the Hubble Constant, Spatial Curvature, and the Dark Energy Equation of State. *Astrophysical Journal*, 2010, v. 711, 201–221.
27. Tytler D. et al. Cosmological Parameters σ_8, the baryon density Ω_b, the vacuum energy density Ω_Λ, the Hubble constant and the UV background Intensity from a calibrated measurement of H I Ly? absorption at $z = 1.9$. *Astrophysical Journal*, 2004, v. 617, 1–28.
28. Weinberg S. Cosmology, Oxford Univ. Press, 2008
29. Corasaniti P.S., Giannantonio T., Melchiorri A. Constraining dark energy with cross-correlated CMB and large scale structure data. *Physical Review D*, 2005, v. 71, 123521.
30. Sherwin B.D. et al. Evidence for Dark Energy from the Cosmic Microwave Background Alone Using the Atacama Cosmology Telescope Lensing Measurements. *Physical Review Letters*, 2011, v. 107, 021302.
31. Sawangwit U. Shanks T. Beam profile sensitivity of the WMAP CMB power spectrum. *Monthly Notices of the Royal Astronomical Society*, 2010a, v. 407, L16–L20.
32. Sawangwit U., Shanks T. Lambda-CDM and the WMAP power spectrum beam profile sensitivity. 2010, eprint arXiv:1006.1270.

ΔI=1 Signature Splitting in Signature Partners of Odd Mass Superdeformed Nuclei

A.M. Khalaf*, M.D. Okash†, M.H. Ghomiem* and W.A. Muhammad*

*Physics Department, Faculty of Science, Al-Azhar University, Cairo, Egypt. E-mail: ali_khalaf43@hotmail.com
†Physics Department, Faculty of Science (Girls College), Al-Azhar University, Cairo, Egypt

The spins, transition energies, rotational frequencies, kinematic and dynamic moment of inertia of rotational bands of signature partners pairs of odd–A superdeformed bands in A∼190 region were calculated by proposing a simple model based on collective rotational model. Simulated search program was written to determine the model parameters. The calculated results agree with experimental data for fourteen signature partner pairs in Hg/Tl/Pb/Bi/nuclei. We investigated the ΔI=1 signature splitting by extracted the difference between the average transitions I+2 ⟶ I and I ⟶ I-2 energies in one band and the transition I+1 ⟶ I-1 energies in its signature partner. Most of the signature partners in this region show large amplitude staggering. The signature splitting has the effect of increasing dynamical moment of inertia J^2 for favored band and decreasing J^2 for the unfavored band.

1 Introduction

Since the first observation of superdeformation in ^{152}Dy [1] and in ^{191}Hg [2] more than 350 settled SD bands in more than 100 nuclei have been will established in several mass regions of nuclear chart A∼190, 150, 130 [3–6]. With the aid of large γ-ray detectors arrays, new regions of SD nuclei have been discovered encircle mass A∼80, 60, 70, 90 regions. The A∼190 mass region is of special interest, more than 85 SD bands have now been observed in this mass region alone in Au, Hg, Tl, Pb, Bi and Po nuclei. The SD states in A∼190 mass region have been observed down to quite low spin also many SD bands in the A∼190 show the same smooth rise in the dynamical moment of inertia as rotational frequency increase, which is associated [7–9] with the successive gradual alignments of a pair of nucleons occupying specific high-N intruder orbitals in the presence of pairing correlations.

Spin assignment is one of the most difficult and still unsolved problems in the study of nuclear superdeformation, because spins have not been determined experimentally in SD nuclei. This is due to the difficulty of establishing the excitation of a SD band into known yrast states. Several related approaches to assign the spins of SD bands in terms of there observed γ-ray energies were proposed [10–28]. For all approaches an extrapolation fitting procedure was used.

The development of large γ-ray arrays has allowed experimentalists to find new phenomena at high angular momenta. For example some SD nuclear bands in mass regions A∼150 and A∼190 show an unexpected regular staggering effects in the transition energies E_γ (a zigzag behavior as a function of rotational frequency or spin). At high rotational frequencies a ΔI=2 staggering was observed [29–39]. It has attracted much attention and interest, and has thus become one of the most frequently debated subjects.

The ΔI=2 rotational bands are perturbed and two ΔI=4 rotational sequences emerge with an energy splitting of about some hundred eV. This is commonly called ΔI=4 bifurcation or as C_4 oscillation, because the SD-energy levels are consequently separated into two spin sequences with spin values I_0, I_0+4, I_0+8,... and I_0+2, I_0+6, I_0+10,... respectively.

Many ΔI=1 signature splitting have been observed in ND nuclei for different bands, like odd-even staggering (OES) in the gamma vibrational band at low spin [40], the beat odd-even ΔI=1 staggering patterns observed in the octuple bands [41] and the ΔI=1 odd-even staggering structure of alternating parity bands in even-even nuclei [42, 43].

There is another kind of staggering happens in SD odd-A nuclei, the ΔI=1 signature splitting in signature partner pairs. It was seen that most of SD rotational bands in odd-A nuclei in the A∼190 region are signature partners [44–53]. Most of these signature partners show large amplitude signature splitting and the bandhead moments of inertia of each pair are almost identical.

2 Sketch of the Model

In the model used, the excitation energy of a SD State E(I) and spin I is expressed as:

$$\hat{I}^2 = I\,(I+1) = \sum_n b_n\, E^n(I). \tag{1}$$

With $\hat{I}\ [I\,(I+1)]^{1/2}$. If we restrict to three terms only, then

$$I(I+1) = b_0 + b_1 E(I) + b_2 E^2(I). \tag{2}$$

Solving for E(I) we get the two-parameters formula for E(I)

$$E(I) = E_0 + a\left([1 + bI(I+1)]^{1/2}\right) \tag{3}$$

with a,b and E_0 simply expressed by b_0, b_1 and b_2

$$a = \frac{1}{2b_2}[b_1^2 - 4b_0 b_2]^{1/2} \tag{4}$$

$$b = \frac{4b_2}{b_1^2 - 4b_0b_2} \tag{5}$$

$$E_0 = a - \frac{b_1}{2b_2} \tag{6}$$

where b characterizes the nuclear softness.

The rigid rotor limit corresponds to b$\longrightarrow$0 and a, E_0 keeping finite. The value of the parameter a increases slowly with I. It is expected that a better expression may be obtained if a weak I dependence of the parameter a is taken into account. So equation (3) is tentatively modified as follows:

$$E(I) = a\Big([1 + b(I)(I+1)]^{1/2} - 1\Big) + cI(I+1) \tag{7}$$

with an additional parameter c. Leading to a form for the gamma transition energies

$$\begin{aligned} E_\gamma(I) &= a\Big([1+bI(I+1)]^{1/2} - [1+b(I-2)(I-1)]^{1/2}\Big) \\ &\quad +2c(2I+1). \end{aligned} \tag{8}$$

The kinematic J^1 and dynamic J^2 moment of inertia associated with the a, b, c, formula are:

$$J^1 = ab[1 + bI(I+1)]^{1/2} + \frac{1}{2c} \tag{9}$$

$$J^2 = ab[1 + bI(I+1)]^{3/2} + \frac{1}{2c}. \tag{10}$$

The bandhead moment of inertia is

$$J_0 = \frac{\hbar^2}{ab + 2c}.$$

Each SD nucleus is described by three adjustable parameters a, b and c which are determined by fitting procedure of all known levels.

For the SD bands, one can extract the rotational frequency, dynamic and kinematic moment of inertia by using the experimental intra band E_2 transition energies as follows:

$$\hbar\omega = \frac{1}{4}[E_\gamma(I+2) + E_\gamma(I)] \tag{11}$$

$$J^2(I) = \frac{4\,\hbar^2}{\triangle E_\gamma} \tag{12}$$

$$J^1(I-1) = \frac{\hbar^2(2I-1)}{E_\gamma} \tag{13}$$

where

$$E_\gamma = E(I) - E(I-2),$$

$$\triangle E_\gamma = E_\gamma(I+2) - E_\gamma(I).$$

It is seen that whereas the extracted J^1 depends on I proposition, J^2 does not.

3 Analysis of ΔI=1 signature splitting in SD signature partner

To investigate the ΔI=1 staggering in signature partner pairs of odd SD band, one must extract the differences between the average transition I+2$\longrightarrow$I and I$\longrightarrow$I-2 energies in one band the transition I+1$\longrightarrow$I and I$\longrightarrow$I-1 energies in the signature partner

$$\begin{aligned} \triangle^2 E_\gamma(I) &= \tfrac{1}{2}\Big[E_\gamma(I+2 \longrightarrow I) + E_\gamma(I \longrightarrow I-2) \\ &\quad -2E_\gamma(I+1 \longrightarrow I-1)\Big] \end{aligned}$$

where E_γ(I) is proposed in equation (8).

4 Numerical Calculation and Discussions

Our selected data set includes fourteen signature partner pairs in ten odd SD nuclei in the A$\sim$ 190 mass region, namely:

^{191}Hg (SD2, SD3)	^{193}Hg (SD1, SD2)	^{193}Hg(SD3, SD4)
^{193}Hg (SD3, SD4)	^{195}Hg (SD3, SD4)	^{191}Tl(SD1, SD2)
^{193}Tl (SD1, SD2)	^{195}Tl (SD1, SD2)	^{193}Pb(SD3, SD4)
^{193}Pb (SD5, SD6)	^{195}Pb (SD1, SD2)	^{195}Pb(SD3, SD4)
^{197}Pb (SD1, SD2)	^{197}Bi (SD2, SD3)	

The experimental transition energies are taken from reference [3]. To parameterize the spins of the SD bands, we assumed various values for the bandhead spin I_0 for each SD band and the model parameters a, b and c are adjusted by using a computer simulated search program in order to obtain a minimum root mean square deviation

$$\chi = \left[\frac{1}{N}\sum_{i=1}^{N}\frac{E_\gamma^{exp}(I) - E_\gamma^{Theor}(I)}{\triangle E_\gamma^{exp}(I)}\right]^{1/2}.$$

Of the calculated energies E_γ^{cal} from the observed energies E_γ^{exp}, where N is the number of data points considered and $\triangle E_\gamma^{exp}$ is the uncertainty of the γ-transition energies. The fitting procedure was repeated with spin I_0 fixed at the nearest half integer.

Table (1) gives the optimized model parameters a, b, c, the bandhead spin proposition I_0 and the lowest transition energies E_γ (I_0+2 $\longrightarrow$ I_0) for each SD band.

The systematic behavior of kinematic J^1 and dynamic J^2 moments of inertia are guideline for the spin prediction and to understand the properties of the SD bands. We studied the variation of $J^1$and J^2as a function of rotational frequency $\hbar\omega$. The value of $J^1$and J^2 approaches each other at the bandhead spin I_0 . The J^1 moment of inertia is found to be smaller than that of J^2 for all values of $\hbar\omega$. Both $J^1$and J^2 plots are concave upwards. In general the bandhead moments of inertia in our selected signature partners odd-A SD nuclei $J_0 \cong (94 \pm 4)\hbar$ MeV^{-1} are longer than that of the yrast SD bands in neighboring even-even nuclei. The best fitted parameters were used to calculate the theoretical transition energies extracted from our proposed model.

Table 1: The calculated best model parameters a, b, c and suggested bandhead spins I_0 for our selected signature partners in the odd SD nuclei in A $\simeq$ 190 region.

SD Bands	a MeV	b 10^{-4} MeV	c MeV	I_0 $\hbar$	E_γ keV
^{191}Hg(SD2)	19074.6639	3.0809	2.3765	10.5	252.4
^{191}Hg(SD3)	15810.8517	3.6987	2.4037	11.5	272
^{193}Hg(SD1)	1569.7883	23.4445	3.7662	9.5	233.2
^{193}Hg(SD2)	12654.6097	4.3858	2.6051	10.5	254
^{193}Hg(SD3)	12243.4329	4.4984	2.6289	9.5	233.5
^{193}Hg(SD4)	12654.6098	4.3858	2.6051	10.5	254
^{195}Hg(SD3)	72779.9405	1.8708	-0.9723	10.5	284.5
^{195}Hg(SD4)	22034.6647	2.4110	2.4673	15.5	341.9
^{191}Tl(SD1)	307519.2819	0.7272	-5.7903	11.5	276.77
^{191}Tl(SD2)	249002.6385	0.8532	-5.2350	12.5	296.75
^{193}Tl(SD1)	13573.6592	3.7666	2.6759	9.5	227.3
^{193}Tl(SD2)	6380.8736	5.3776	3.5196	8.5	206.6
^{195}Tl(SD1)	6380.8738	5.3776	3.5196	5.5	146.2
^{195}Tl(SD2)	33124.3911	2.4266	1.2551	6.5	167.5
^{193}Pb(SD3)	4702.3802	6.2778	3.8243	10.5	251.5
^{193}Pb(SD4)	16892.1756	3.5957	2.2986	11.5	273
^{193}Pb(SD5)	4337.5276	8.2523	3.6196	8.5	213.2
^{193}Pb(SD6)	3574.7877	9.4219	3.7378	9.5	234.6
^{195}Pb(SD1)	600.9413	13.4593	4.6737	7.5	162.58
^{195}Pb(SD2)	15864555.765	0.0139	-5.9659	6.5	182.13
^{195}Pb(SD3)	2362.3559	13.4225	3.9167	7.5	198.2
^{195}Pb(SD4)	18884.3711	3.5500	2.0732	8.5	213.6
^{197}Pb(SD1)	9713.0371	2.5870	3.8497	7.5	183.7
^{197}Pb(SD2)	724986.6813	0.1798	-1.3692	6.5	204.6
^{197}Bi (SD2)	6.09E+08	8.24E-07	-245.7806	8.5	166.2
^{197}Bi (SD3)	6.09E+08	8.24E-07	-245.7806	9.5	186.7

To investigate the ΔI=1 signature splitting, the difference between the averaged transitions I+2⟶I and I⟶I-2 energies in one band and the transition I+1⟶I-1 energies in its signature partner $\triangle^2 E_\gamma$(I) are determined and its value as a function of spin I for each signature partner pairs are plotted in figure (1). Most of there signature partners show large amplitude staggering with the exception of ^{193}Hg (SD1, SD2), ^{193}Pb (SD5, SD6) and ^{195}Pb (SD3, SD4).

A clear out amplification of $\triangle^2$E(I) is seen in ^{193}Pb (SD3, SD4). For most cases one finds that $\triangle^2$E(I) is very small at lower spins, increasing faster and faster as the spin I increase. The of $\triangle^2$E(I) in ^{193}Tl (SD1, SD2) and ^{195}Tl (SD1, SD2) are remarkable similar.

5 Conclusion

The nuclear superdeformed rotational bands of signature partners of odd-mass number in the A∼190 region have been studied in the framework of a simple formula based on collective rotational model containing three parameters. The formula connected directly the unknown spin and the energy of the level the spins of the observed levels were extracted by assuming various values to the lowest spin of the bandhead at the nearest half integer. The optimized three parameters have been deduced by using a computer simulated search program in order to obtain a minimum root mean square deviation of the calculated transition energies from the measured energies.

The calculated transition energies, level, spins, rotational frequencies, kinematic and dynamic moments of inertia are examined for fourteen signature partner pairs. To investigate the ΔI=1 signature splitting for each signature partner pair, we calculated the difference between the average transitions I+2 ⟶I and I⟶I-2 energies in one band and the transition I+1⟶I-1 energies in its signature partner. Most of the signature partners in this region show large amplitude ΔI=1 staggering.

Submitted on April 6, 2013 / Accepted on April 19, 2013

References

1. Twin P.J., Nyak B.M. Observation of a Discrete-Line Superdeformed Band up to 60$\hbar$ in ^{152}Dy. *Physical Review Letters*, 1986, v. 57, 811–814.

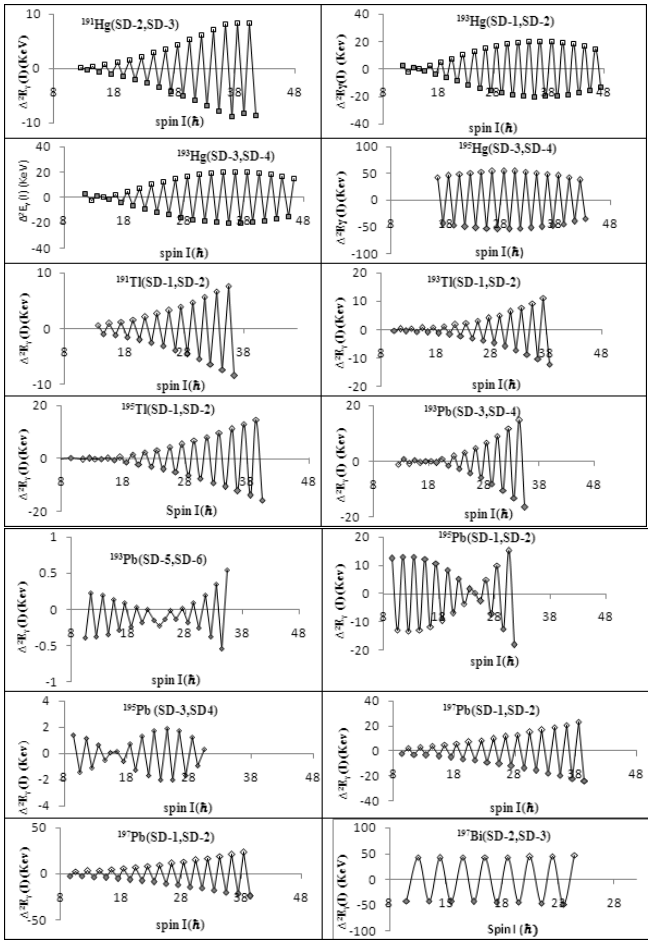

Fig. 1: The ΔI=1 signature splitting in some signature partners of odd-A superdeformed nuclei.

2. Moore E. F., Janssens R.V.F et al. Observation of superdeformation in 191Hg. *Physical Review Letters*, 1989, v. 63, 360–363.
3. Singh Balraj. Table of Superdeformed Nuclear Bands and Fission Isomers. *Nuclear Data Sheets*, 2006, v. 107 1–224.
4. Singh Balraj, Zywine Roy, Richard B. Firestone. Table of Superdeformed Nuclear Bands and Fission Isomers. *Nuclear Data Sheets*, 2002, v. 97 (1), 241–592.
5. Han X.L. and Wu C.L. At. Data NUCLEAR SUPERDEFORMATION DATA TABLES. *Nuclear Data Tables*, 1996, v. 117, 117.
6. Singh B., Firestone R.B., Chu S.Y.E. Table of superdeformed nuclear bands and fission isomers. *Egypt Journal of Physics*, 1996, v. 78, 1–177.
7. Ye D., Janssens R.V.F. et al. Superdeformed band in ^{192}Hg. *Physics Review* , 1990, v. C41, R13–R17.
8. Riley M.A., Cullen D.M. et al. Multiple superdeformed bands in ^{194}Hg and their dynamical moments of inertia. *Nuclear Physics*, 1990, v. A512, 178–188.
9. Drigert M.W., Carpenter M.P. et al. Superdeformed bands in ^{189,190}Hg. *Nuclear Physics*, 1991, v. A530, 452–474.
10. Stephens F.S. Spin alignment in superdeformed rotational bands. *Nuclear Physics* , 1990, v. A520 c91–c104.
11. Becker J.A., N. Roy, Henry E.A. et al. Level spin and moments of inertia in superdeformed nuclei near A=194. *Nuclear Physics*, 1990, v. A520, c187–c194.
12. Draper J.E., Stephens F.S. et al. Spins in superdeformed bands in the mass 190 region 190 region. *Physics Review* , 1990, v. C42, R1791–R1795 .
13. J. E. Draper J.E., Stephens F.S. et al. J Superdeformation in ^{196,198}Pb. *ibid*, 1991, v. 42, R179.
14. Zeng J.Y., Meng J. et al. Spin determination and quantized alignment in the superdeformed bands in ^{152}Dy, ^{151}Tb, and ^{150}Gd. *Physical Review*, 1991, v. C44, R1745–R1748.
15. Becker J.A., Henry E.A. et al. Level spin for superdeformed nuclei near A=194. *Physical Review* , 1992, v. C46, 889–903.
16. Piepenbring R., Protasov K.V. Superfluid liquid model with triplet pairing for superdeformed in A 130-150 region. *Z.Physics* , 1993, v. A347 (7), 27–35.
17. Xu Furong , Hu Jimin et al. Cranking Bohr-Mottelson Hamiltonian applied to superdeformed bands in A 190 region. *Physical Review*, 1994, v. C49, 1449–1453.
18. Fan H-Y., Jing S.-C. Even and Odd Binomial States. *Communications in Theoretical Physics*, 1995, v. 24, 125–128.
19. Hegazi A.M., Ghoniem M.H., Khalaf A.M. Superdeformation in ^{154}Er. *Egypt Journal of Physics*, 1999, v. 30 (3), 293–R1174.
20. Khalaf A.M. et al. Band Head of the Superdeformed Bands in the $A \sim 150$ Mass region Nuclei. *Egypt Journal of Physics*, 2002, v. 33 (1), 67–87.
21. Khalaf A.M. et al. Spin Prediction and Systematics of Moments of inertia of superdeformed Nuclear Rotational Band in the Mass Region A∼190. *Egypt Journal of Physics*, 2002, v. 33 (3), 585–602.
22. Khalaf A.M. et al. Deascription of Rotational Bands in Superdeformed Nuclei by Using Two-parameter Empirical Formula. *Egypt Journal of Physics*, 2003, v. 34 (2), 159–177.
23. Khalaf A.M. et al. Properties of Superdeformed Rotational Bands of Odd Nuclei in the Mass-190 Region Using Harris Expansion. *Egypt Journal of Physics*, 2003, v. 34 (2), 195–215.
24. Khalaf A.M. et al. Analysis of Rotational Bands in Superdeformed Nuclei Using sdg Interacting Boson Model. *Egypt Journal of Physics*, 2004, v. 34 (1), 79–104.
25. Khalaf A.M. and Sirag M.M. Prediction of Nuclear Superdeformed Rotational Bands Using Incremental Alignments. *Egypt Journal of Physics*, 2006, v. 37 (3), 277–293.
26. Khalaf A.M., Allam M.A., Saber E. Rotational Bands of Superdeformed Nuclei in Framework of Variable Moment of Inertia Model. *Egypt Jurnal of Physics*, 2006, v. 73 (3), 195.
27. Khalaf A.M., Allam M.A., Saber E. Signature Partners in Odd Superdeformed Nuclei in Mass Region A∼190. *Egypt Journal of Physics*, 2008, v. 39 (1), 41–65.
28. Khalaf A.M., Allam M.A. and Sirag M.M. Bandhead Spin Determination and Moments of inertia of Superdeformed Nuclei in Mass Region 60-90 Using Variable Moment of inertia Model. *Egypt Journal of Physics*, 2010, v. 41 (2), 13–27.
29. Flibotte S. et al. $\Delta I = 4$ bifurcation in a superdeformed band: Evidence for a C_4 symmetry band. *Physical Review Letters*, 1993, v. 71, 4299–4302.
30. Cederwall B. et al. New features of superdeformed bands in ^{194}Hg. *Physica Scripta*, 1994, v. 72 3150–3153.
31. Flibotte S. Hackman G. et al. Multi-particle excitations in the superdeformed ^{149}Gd nucleus. *Nuclear Physics*, 1995, v. A584, 373–396.
32. Carpenter M.P., Janssens R.V.F. Identification of the unfavored N=7 superdeformed band in ^{191}Hg. *Physical Review* , 1995, v. 51, 2400–2405.
33. Bernstein L.A. and Hughes J.R. Superdeformation in ^{154}Er. *Physical Review*, 1995, v. C52, R1171–R1174.

34. de Angelis G. and Wyss R. Spectroscopy in the second well of the ^{148}Gd nucleus Two quasiparticle and collective excitations. *Physical Review*, 1996, v. C35, 679–688.
35. Fischer S.M., Carpenter M.P. et al. Alignment additivity in the two-quasiparticle superdeformed bands of ^{192}Tl. *Physical Review*, 1996, v. C35, 2126—2133.
36. Semple A.T., Nolan P.J. Energy Staggering in Superdeformed bands in ^{131}Ce, ^{132}Ce and ^{133}Ce. *Physical Review Letter*, 1996, v. 76, 3671–3674.
37. Krücken R., Hackman G. et al. Test of ΔI=2 staggering in the superdeformed bands of ^{194}Hg. *Physical Review*, 1996, v. C45, R2109–R2113.
38. Cederwall B. et al. PROPERTIES OF SUPERDEFORMED BANDS IN DY-153. *Physical Review*, 1995, v. B346, 244–250.
39. Haslip D.S., Flibotte S., and de France G. Δ*I*=4 Bifurcation in Identical Superdeformed Bands. *Physical Review Letters*, 1997, v. 78, 3447–3450.
40. Singh M., Bihari C. et al. Evidence of rigid triaxiality in some xenon nuclei. *Canadian Journal of Physics*, 1995, v. 85 (8), 899–910.
41. Bonatsos D., Daskaloyannis C. Δ*I*=1 staggering in octupole bands of light actinides: "Beat" patterns. *Physics Letters*, 2000, v. C62, 24301–24313.
42. Wiedenhöver I, Janssens R.V.F. et al. Octupole Correlations in the Pu Isotopes: From Vibration to Static Deformation. *Physical Review Letters*, 1999, v. 83, 2143–2146.
43. Minkov N., Yotov P., Drenska S. Parity shift and beat staggering structure of octupole bands in a collective model for quadrupole-octupole deformed nuclei. *Journal of Physics G – Nuclear Physics*, 2006, v. 32, 497–503.
44. Duprat J. Azaiez F. et al. M1 transitions between superdeformed states in ^{195}Tl: the fingerprint of the i132 proton intruder orbital. *Physical Letter* , 1994, v. B341, 6–11.
45. Joyce M.J. et al. First measurement of magnetic properties in a superdeformed nucleus:^{193}Hg. *Physical Review Letters*, 1993, v. 71, 2176–2179.
46. Hughes J.R., Becker J.R. et al. Superdeformation in ^{193}Pb and the effects of the N=7 intruder orbital. *Physical Review* , 1990, v. C51, R447–R451.
47. Farris L.P., Henry E.A. et al. Neutron blocking and delayed proton pair alignment in superdeformed ^{195}Pb. *Physical Review*, 1996, v. C 51, R2288–R2292.
48. Baunem S. et al. The $i_{13/2}$ proton intruder orbital in the superdeformed ^{193}Tl nucleus: Effective magnetic moment and blocking of proton pairing. *Physical Review*, 1996, v. C53, R9–R13.
49. Carpenter M.P., Janssens R.V.F. and Flocard H. Identification of the unfavored N=7 superdeformed band in ^{191}Hg. *Physical Review*, 1995, v. C51, 2400–2405.
50. Joyce M.J., Sharpey-Schafer J.V. and Flocard H. The N = 7 unfavoured superdeformed band in ^{193}Hg; coriolis splitting and neutron shell structure at extreme deformation *Physical Letter*, 1999, v. B340, 150–154.
51. Joyce M.J., Sharpey-Schafer and Riley M.A. Microscopic study of the properties of identical bands in the A=150 mass region *Physical Review*, 1999, v. C59, 3120–3127.
52. Hackman G. and Krücken R. Structure of superdeformed bands in ^{195}Hg *Physical Review*, 1997, v. C55, 148–154.
53. Clark R.M., Bouneau S. et al. Superdeformation in the bismuth nuclei. *Physical Review*, 1995, v. C51, R1052–R1056.
54. Hibbert I.M., Wadsworth R. et al. Superdeformed structures in ^{197,198}Pb. *Physical Review*, 1996, v. C54, 2253–2258.

Nuclear Shape Transition Using Interacting Boson Model with the Intrinsic Coherent State

A.M. Khalaf*, H.S. Hamdy† and M.M. El Sawy‡

*Physics Department, Faculty of Science, Al-Azhar University, Cairo, Egypt. E-mail: ali_khalaf43@hotmail.com
†Department of Physics, Faculty of Science, Beni-Suef University, Beni-Suef, Egypt.
‡Demonstrator at Basic Science department, Engineering Faculty, The British University, Cairo, Egypt. E-mail: maiamy4@gmail.com

The values of the potential energy surface (PES) for the even-even isotopic chains of Nd/Sm/Gd/Dy are studied systematically using the simplified form of interacting boson model (IBM) with intrinsic coherent state. The critical points have been determined for each isotope chain. The phase diagrams exhibits first-order shape phase transition from spherical U(5) to deformed axial symmetric prolate SU(3) when moving from light isotopes to heavy ones.

1 Introduction

We note that in the interacting boson model-1 (IBM-1) [1, 2] one describes an even-even nucleus as a system of N bosons able to occupy two levels, one with angular momentum restricted to zero (s boson) and one with angular momentum 2 (d boson).

The bosons are assumed to interact via a two-body residual interaction. Denoting by b_i (i=1,...,6) the creation (annihilation) operators for bosons ($b_1 = s, b_{2,...,6} = d$) it is easy to see that the 36 operators $G_{ii'} = b_i^\dagger b_{i'}$ close under the Lie algebra of U(6). This simple model allows the utilization of algebraic symmetric for approaching different type of nuclear spectra, known as dynamical symmetries and corresponding to un-harmonic vibrator (U(5) Symmetry) [3], rigid deformations (SU(3) Symmetry) [4] and γ-instability (O(6) Symmetry) [5]. In these special cases it is possible to find analytical solutions of the boson Hamiltonian and deal with small deviations from these symmetries using different perturbation methods.

However, real nuclei may deviate considerably from the simple dynamical limits. This is represented in the Casten triangle [1–6] with vertices corresponding to the standard dynamical symmetries and the sides of the triangle represent direct transition between the limiting cases, whereas all complex transition regions are contained in the area. Phase transitions between these shapes were studied, and it is known that the phase transition from U(5) to O(6) is second order, while any other transitions within the Casten triangle from a spherical to deformed shape is first order [7–23].

Now, there is a class of symmetries that are formulated in terms of the Bohr Hamiltonian and that can be applied to critical point situation [24–26]. In particular, at the critical point from spherical to γ-unstable shapes, called E(5) [24], at the critical point from spherical to axially deformed shapes, called X(5) [25] and the critical point from axially deformed shapes to triaxial shapes, called Y(5) [26]. Since the introduction of these limits many theoretical [27–32] and experimental [33–39] studies have been presented in order to look for nuclei that exhibit the properties of critically and to classify the corresponding phase transitions. Many studies have extended these original models to more complex situations [40–44].

The relation between the Bohr-Mottelson collective model [45] and the IBM was established [46, 47] on the basis of an intrinsic (or coherent) state for the IBM. Via this coherent state formalism, a potential energy surface (PES) E(β, γ) in the quadruple deformation variables β and γ can be derived for any IBM Hamiltonian and the equilibrium deformation parameters β_0 and γ_0 are then found by minimizing E(β, γ). The deformation parameter β measures the axial deviation from sphericity, while the angle variable γ controls the departure from axial symmetry.

In the present work, we investigate shape phase transition within the IBM-1 using coherent state formalism for various rare earth isotopic chains. The paper is organized as follows. First the IBM and the symmetry triangle used in the present work is briefly described in section 2. In this variation of the IBM, the coherent state approach is treated to produce PES's in section 3. The location of the critical point in the shape transition is identified in section 4. We review the concept of dynamical symmetry in section 5. In section 6 a systematic study of isotopic chains on Nd/Sm/Gd/Dy related to the U(5)-SU(3) shape transition is given and main conclusions arising from the present results are discussed.

2 The IBM-1 Hamiltonian and Coherent State

Denoting by $C_n[G]$ the n^{th}-order Casmir operator of the Lie group G, the general sd-IBM Hamiltonian with up to two-body interactions can be written in the following form:

$$\begin{aligned} H &= \epsilon C_1[U(5)] + k_1 C_2[U(5)] \\ &+ k_2 C_2[O(5)] + k_3 C_2[O(3)] \\ &+ k_4 C_2[SU(3)] + k_5[O(6)] \end{aligned} \quad (1)$$

The Casmir operators are defined by the following equations

$$C_1[U(5)] = \hat{n}_d \quad (2)$$

$$C_2[U(5)] = \hat{n}_d(\hat{n}_d + 4) \tag{3}$$

$$C_2[O(5)] = 4\left[\frac{1}{10}(\hat{L}.\hat{L} + \hat{T}_3.\hat{T}_3)\right] \tag{4}$$

$$C_2[O(3)] = 2(\hat{L}.\hat{L}) \tag{5}$$

$$C_2[SU(3)] = \frac{2}{3}\left[2(\hat{Q}.\hat{Q}) + \frac{3}{4}(\hat{L}.\hat{L})\right] \tag{6}$$

$$C_2[O(6)] = 2\left[N(N+4) - 4(\hat{P}.\hat{P})\right] \tag{7}$$

where $\hat{n}_d$, $\hat{P}$, $\hat{L}$, $\hat{Q}$ and $\hat{T}_3$ are the boson number, pairing, angular momentum, quadruple and octuple operators defined as

$$\hat{n}_d = (d^\dagger \tilde{d})^{(0)} \tag{8}$$

$$\hat{P} = \frac{1}{2}(\tilde{d}\tilde{d}) - \frac{1}{2}(\tilde{s}\tilde{s}) \tag{9}$$

$$\hat{Q}[\chi] = [d^\dagger \tilde{s} + s^\dagger \tilde{d}]^2 + \chi[d^\dagger x \tilde{d}]^{(2)} \tag{10}$$

$$\hat{L} = \sqrt{10}[d^\dagger x \tilde{d}]^{(1)} \tag{11}$$

$$\hat{T}_3 = [d^\dagger x \tilde{d}]^{(3)} \tag{12}$$

where $s^\dagger(s)$ and $d^\dagger(d)$ are monopole and quadruple boson creation (annihilation) operators respectively. The study of shape phase transition in even-even nuclei can be well done from the simple two parameter IBM Hamiltonian, the well known consistent-Q Hamiltonian

$$H = \varepsilon \hat{n}_d - k\hat{Q}(\chi) \cdot \hat{Q}(\chi). \tag{13}$$

The symbol $(\cdot)$ represents the scalar product and the scaler product of two operators with angular momentum L is defined as $\hat{T}_L.\hat{T}_L = \Sigma_M(-1)^M \hat{T}_{LM}\hat{T}_{L-M}$ where $\hat{T}_{LM}$ corresponds to the M component of the operator $\hat{T}_L$.

The Hamiltonian of equation (13) describes the main features of collective nuclei, it contains the dynamical symmetries of the IBM for spherical choices of the coefficients ε, k and χ, and allows to describe the transitional regions between any of symmetry limits as well. In discussing phase transitions, it is convenient to introduce the control parameter η, such as:

$$\frac{\eta}{1-\eta} = \frac{1}{N}\frac{\varepsilon}{k} \tag{14}$$

where N is the total number of boson. Hamiltonian(1) can be written in the second form

$$H = C\left[\eta\hat{n}_d - \frac{1-\eta}{N}\hat{Q}(\chi).\hat{Q}(\chi)\right]. \tag{15}$$

With

$$C = \varepsilon + Nk, \qquad \eta = \frac{\varepsilon/k}{N + \varepsilon/k}. \tag{16}$$

The second form equation (15) avoids the infinities inherent in the use of the ratio of ε/k as η varies from 0 to 1. The factor C in equation (15) is only a scale factor and η and χ are therefore the two parameters that determine the structure. The values of the control parameter η ranges from 0 to 1 and χ is located in the interval of $-\sqrt{7}/2$ (-1.32) to $\sqrt{7}/2$ (+1.32).

Let us consider the Hamiltonian of equation (5) and the effects of its two parameters η and χ. Clearly, one of the most important features of the IBM is the existence of three distinct dynamical symmetries (DS), each representing a well defined phase of nuclear collective motion. The three DS are: the U(5) symmetry for spherical vibrational nuclei (η=1), the SU(3) symmetry for prolate deformed nuclei (η=0, χ= $-\sqrt{7}/2$) and the O(6) symmetry for γ-unstable deformed nuclei (η=0, χ=0), the SU(3) symmetry for oblate deformed nuclei corresponding to (η=0, χ=+ $\sqrt{7}/2$). For intermediate values of the control parameters η and χ, the potential energy surface (PES) function will describe a certain point on the IBM symmetry triangle located between the three limits.

Comparing the simplified Hamiltonian equation (15) with equation (1) we see that only two terms of the general form are considered. Rewriting equation (15) in the form of equation (1), we get:

$$\begin{aligned} H \;=\; & \left[\eta + \frac{2}{7N}(1-\eta)\chi\left(\chi + \frac{\sqrt{7}}{2}\right)\right]C_1[U(5)] \\ & + \frac{2}{7N}(1-\eta)\chi\left(\chi + \frac{\sqrt{7}}{2}\right)C_2[U(5)] \\ & + \frac{1}{N}(\eta-1)\left(1 + \frac{3}{\sqrt{7}}\chi + \frac{2}{7}\chi^2\right)C_2[O(5)] \\ & + \frac{1}{14N}(1-\eta)\chi\left(\chi + 2\sqrt{7}\right)C_2[0(3)] \\ & + \frac{1}{\sqrt{7}N}(\eta-1)\chi C_2[SU(3)] \\ & + \frac{1}{N}(1-\eta)\left(1 + \frac{2}{\sqrt{7}\chi}\right)C_2[O(6)]. \end{aligned} \tag{17}$$

In IBM-1, the intrinsic coherent normalized state of a nucleus with N valence bosons outside the doubly-closed shell state is given by:

$$|N\beta\gamma\rangle = \frac{1}{\sqrt{N!}}(\Gamma_C^\dagger)^N|0\rangle \tag{18}$$

where $|0\rangle$ denotes the boson vacuum, and

$$\Gamma_C^\dagger = \frac{1}{\sqrt{1+\beta^2}}\left[s^\dagger + \beta\cos\gamma d_0^\dagger + \frac{1}{\sqrt{2}}\beta sin\gamma(d_2^\dagger + d_{-2}^\dagger)\right]. \tag{19}$$

Here $\beta \geq 0$ and $0 \leq \gamma \leq \pi/3$ are intrinsic shape parameters. We get the PES by calculating the expectation value of Hamiltonian (17) on the boson condensate equation (18). The

corresponding PES as a function of the deformations β and γ is given by:

$$\begin{aligned} E(N,\eta,\chi,\beta,\gamma) = \\ = \; & -5(1-\eta)+\frac{1}{(1+\beta^2)^2} \\ & \Big\{\Big[N\eta-(1-\eta)\big(4N+\chi^2-8\big)\Big]\beta^2 \\ & +\left[N\eta-(1-\eta)\left(\frac{(2N+5)}{7}\chi^2-4\right)\right]\beta^4 \\ & +4N(1-\eta)\sqrt{\tfrac{2}{7}}\chi\beta^3\cos 3\gamma\Big\} \end{aligned} \tag{20}$$

3 Location of the Critical Symmetries

Minimization of the PES equation (20) with respect to β for given values of the control parameters η and χ, gives the equilibrium value β_e. The phase transition is signaled by the condition at $\beta = 0$

$$\frac{d^2E}{d\beta^2} = 0, \tag{21}$$

which fixes the critical value of the control parameter η. The critical point in the above equation (20) is given by the value of η where the coefficient at β^2 vanishes, *i.e.*

$$\eta_{critical} = \frac{4N+\chi^2-8}{5N+\chi^2-8}. \tag{22}$$

At this value, the second β derivative for $\beta = 0$ changes its sign, which means that $\beta = 0$ maximum becomes a local minimum. Note that the critical point (22) depends on χ, it changes between: $\eta(-\sqrt{7}/2) = (16N-25)/(20N-25)$ at U(S)-SU(3) side if the symmetry triangle, and $\eta(0) = (16N-32)/(20N-32)$ at the U(5)-O(6) side, condition (12) gives in the case of large-N limit the value 4/5.

If we ignore the contribution of one-body term of the quadruple-quadruple interaction and in large N limit (N-1≃N) and $\gamma = 0$, equation (20) takes the form

$$\begin{aligned} E(N,\eta,\chi,\beta) = \; & \frac{N\beta^2}{(1+\beta^2)^2}\Bigg[5\eta-4+4\sqrt{\frac{2}{7}}\beta\chi(1-\eta) \\ & +\beta^2\left(\eta-\frac{2}{7}\chi^2(1-\eta)\right)\Bigg]. \end{aligned} \tag{23}$$

The deformation parameter $\beta = 0$ is always a stationary point. For $\eta < 4/5$, $\beta = 0$ is a maximum, while for $\eta > 4/5$, it becomes a minimum. In the case of $\eta = 4/5$, $\beta = 0$ is an inflection point. The $\eta = 4/5$ is the point at which a minimum at $\beta = 0$ starts to develop and defines the antispinodal line. For $\chi \neq 0$, there exists a region, where two minima, one spherical and one deformed, coexist. This region is defined by the point at which the $\beta = 0$ minimum appears (antispinodal point) and the point at which the $\beta \neq 0$ minimum appears (spinodal point). For $\eta = 1$, the system is in the symmetry phase since the PES has a unique minimum at $\beta = 0$. When η decreases, one reaches the spinodal point $\eta = 0.820361$ for $\chi = -\sqrt{7}/2$ as illustrated in Fig. (1) for boson number N=10.

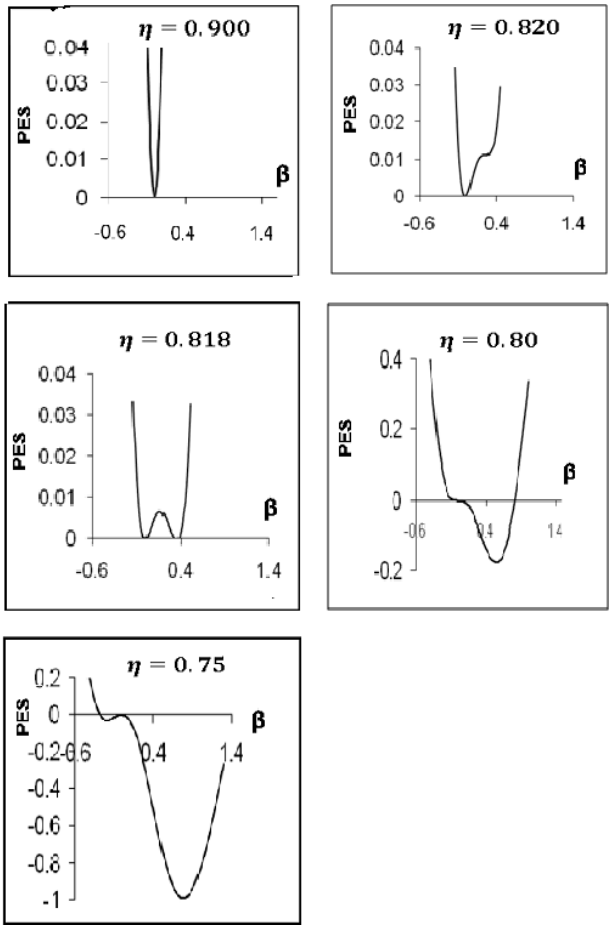

Fig. 1: Potential energy surface (PES) equation (3) for N=10 calculated with IBM without normalization along the axial trajectory $\gamma = 0°, 60°$ as a function of the shape parameter β. The curves describe the first order shape phase transition between spherical to prolate deformed U(5)-SU(3) for control parameter η: $\eta = 0.900$, $\eta = 0.820$ (spinodal), $\eta = 0.818$ (critical point), $\eta = 0.800$ (antispinodal) and $\eta = 0.750$.

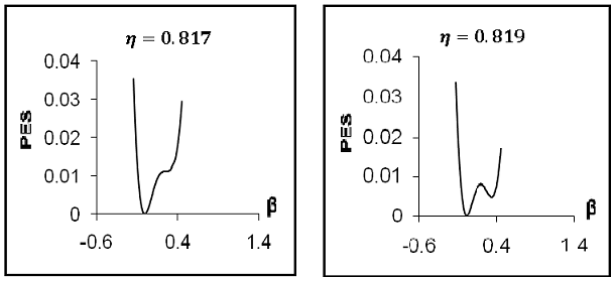

Fig. 2: For two cases in the coexistence region $\eta = 0.817$ and $\eta = 0.819$.

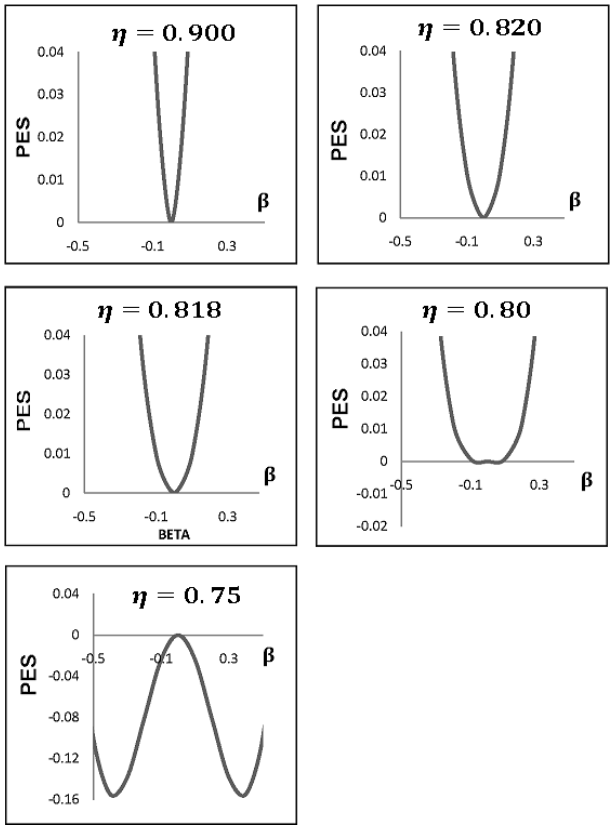

Fig. 3: For $\chi = 0$.

In the coexistence region, the critical point is at the situation in which both minima of spherical point is at the situation in which both minima of spherical and deformed are degenerate. At the critical point, the two degenerated minima are at $\beta_0 = 0$ and $\beta_0 = -\sqrt{7}/4$ and their energy is equal to zero. The critical point line is at $\eta_c = (4 + 2/7\chi^2)/(5 + 2/7\chi^2)$.

The $\chi = -\sqrt{7}/2$ provides $\eta_c = 9/11$ (0.818181). According to the previous analysis, a first order phase transition appears for $\eta \neq 0$, $\chi \neq 0$, while for $\chi = 0$ there is an isolated point of second order phase transition as a function of η. Spinodal, antispinodal and critical point coincide at the critical value $\eta = 4/5$.

We show in Figures (1,2,3) a sketch at this evolution for the special case $\chi = -\sqrt{7}/2$, the two cases in the coexistence region and for $\chi = 0$. From Figure (3), we observe the evolution from the spherical potential $\eta = 0.9$, whose minima is found at $\beta = 0$ to potentials with well-deformed minima $\eta = 0.75$. For intermediate η values one finds a set of potential energy curves which are practically degenerated along the prolate axis in the interval [0, 0.4]. These curves show two minima, on spherical and a prolate deformed one. In particular, for $\eta = 0.81818$, the spherical and the prolate deformed minima are degenerate and this condition defines precisely the critical point of the first order phase transition where the order parameter is the deformation β.

For $\eta = 1$,the Hamiltonian H of equation (15) reduces to the U(5) limit of the IBM corresponds to a spherical shape with vibration

$$H(U(S)) = \hat{n}_d. \tag{24}$$

The PES of H is given by:

$$E(U(5)) = \frac{N\beta^2}{1+\beta^2}. \tag{25}$$

The equilibrium value of the deformation parameter β is easily obtained by solving $\partial E/\partial\beta = 0$ to give $\beta_e = 0$ which corresponds to a spherical shape.

For $\eta = 0$ and $\chi = \mp\sqrt{7}/2$, the schematic Hamiltonian of equation (15) reproduces the SU(3) Limit corresponds to a shape of ellipsoid with rotation (or axial rotation)

$$H((SU(3)) = -\frac{1}{N}\hat{Q}(\chi).\hat{Q}(\chi). \tag{26}$$

If we eliminate the contributions of the one-body terms of quadruple-quadruple interaction, for this case the PES of H is given by:

$$E((SU(3)) = -\frac{(N-1)}{(1+\beta^2)^2}(4\beta^2 + \frac{1}{2}\beta^4 \pm 2\sqrt{2}\beta^3 cos3\gamma). \tag{27}$$

The equilibrium values are given by solving $\frac{dE}{d\beta} = \frac{\partial E}{\partial\gamma} = 0$ to give $\beta_e = \sqrt{2}$ and $\gamma_e = 0$ for $\chi = -\sqrt{7}/2$ and by $\beta_e = \sqrt{2}$ and $\gamma_e = \pi/3$ for $\chi = \sqrt{7}/2$ corresponding to prolate and oblate deformed shape respectively.

For $\eta = 0$ and $\chi = 0$, one recovers the O(6) limit corresponds to γ-unstable

$$H(O(6)) = -\frac{1}{N}\hat{Q}(\chi = 0).\hat{Q}(\chi = 0). \tag{28}$$

Eliminating the one-body terms, the PES depends only on β

$$E(O(6)) = -\frac{(N-1)}{(1+\beta^2)^2}4\beta^2. \tag{29}$$

The equilibrium value is given by $\beta_e = 1$, corresponding to a γ-unstable deformed shape. For intermediate values of the control parameters η and χ, the PES function will describe a certain point on the IBM symmetry triangle, located between the three limits.

4 First-Order U(5)-SU(3) Phase Transition in Nd/Sm/Gd/Dy Rare Earth Nuclei

In a first order phase transition, the state of the rearrangement happens, which means that there involves an irregularity at the critical point.

The study is carried out considering specific isotopic chains of even-even rare earth nuclei $_{60}$Nd, $_{62}$Sm, $_{64}$Gd and $_{66}$Dy displaying first order phase transition from sphericity to

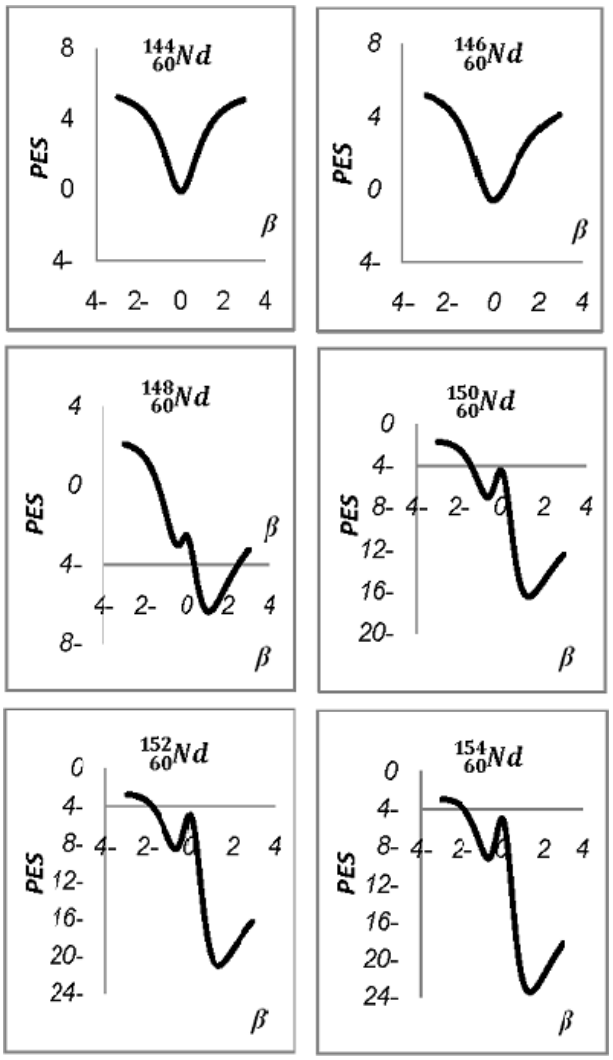

Fig. 4: PES for first order shape phase transition between spherical to prolate deformed U(5)-SU(3) for Neodymium isotope chain $^{144-154}_{60}$Nd (with $N_\pi = 5$ proton bosons and $N_\nu = 1-6$ neutron bosons).

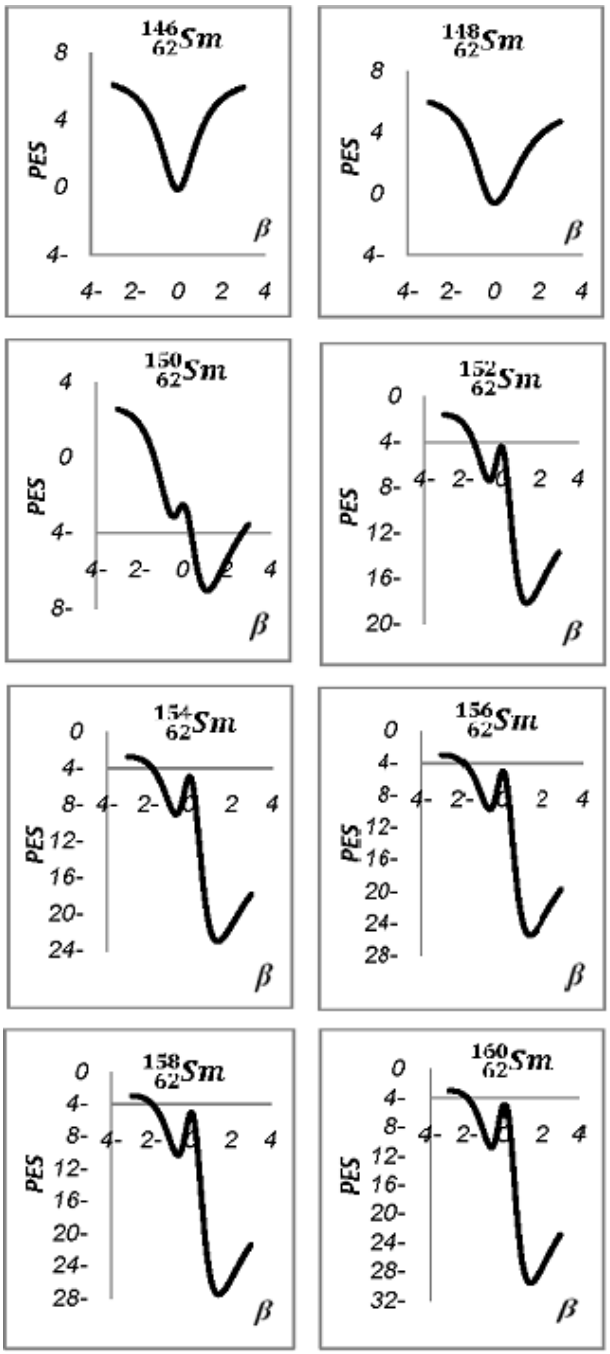

Fig. 5: The same as Fig. (4) but for Samarium isotope chain $^{146-160}_{62}$Sm (with $N_\pi = 6$ proton bosons and $N_\gamma = 1$–8 neutron bosons).

axial symmetric deformed U(S)-SU(3). That is for the nuclei included in this study; all chains begin as vibrational with energy ratio $R_{(4/2)} = E(4_1^+)/E(2_1^+)$ near 2.0 and move towards rotational $R_{(4/2)} = 3.33$ as neutron number is increased. For control parameter $\eta = 1$, we get the U(5) limit and for $\eta = 0$ and $\chi = -\sqrt{7}/2$ the SU(3) limit. For intermediate values of the control parameters η and χ, the PES function will describe a certain point on the IBM symmetry triangle, located between the U(5) and SU(3) limits. To describe a phase transition, one has to establish the values of the control parameter for each nucleus.

For our rare- earth nuclei, we keep χ at the fixed value $\chi = -\sqrt{7}/2$, because some Gd isotopes clearly exhibit the character of the SU(3) dynamical symmetry. This assumption is very successful in describing the Sm nuclei which form neighboring nuclei.

The system passes from the U(5) to the SU(3) limit when the number of bosons is increasing from N=6 towards N=17. The values of the control parameter η is adjusted for each nucleus by using a computer simulated search program in order

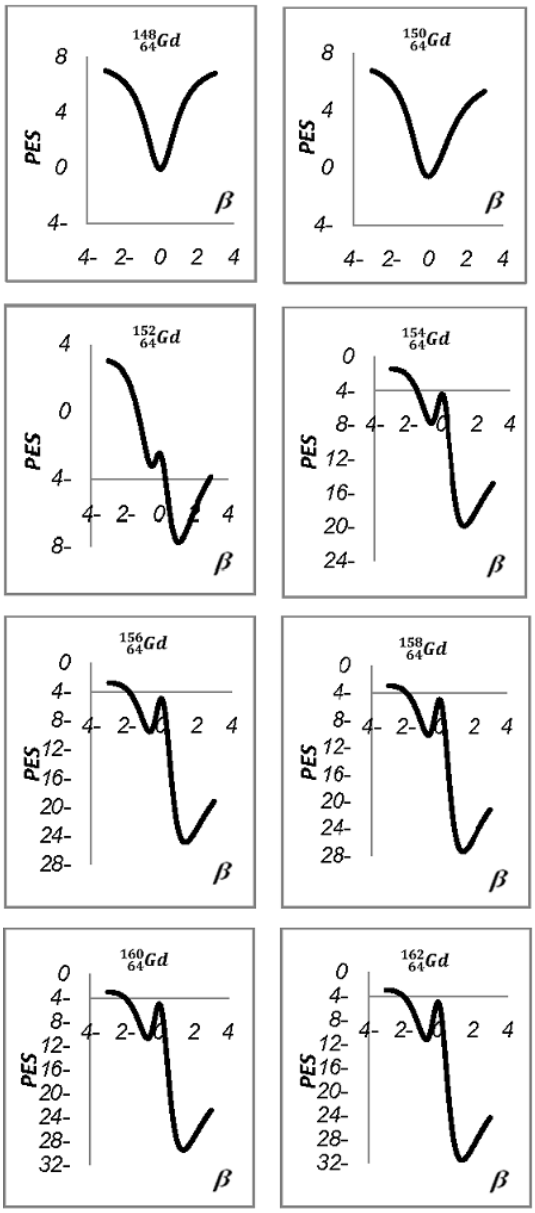

Fig. 6: The same as Fig. (4) but for Gadolinium isotope chain $^{148-162}_{64}$Gd (with N_π = 7 proton bosons and N_γ = 1–9 neutron bosons).

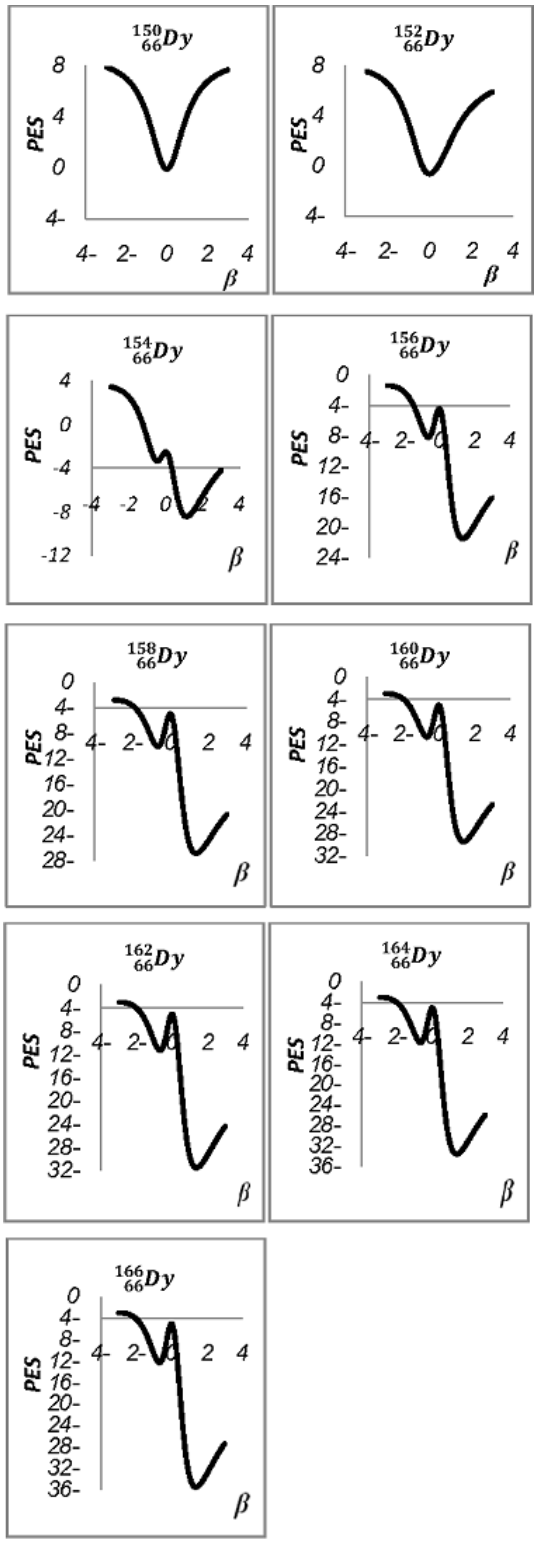

Fig. 7: The same as Fig. (4) but for Dysprosium isotope chain $^{150-166}_{66}$Dy (with N_π = 8 proton bosons and N_γ = 1–9 neutron bosons).

to describe the gradual change in the structure as boson number is varied and to reproduce the properties of the selected states of positive parity excitation ($2^+_1, 4^+_1, 6^+_1, 8^+_1, 0^+_2, 2^+_3, 4^+_3, 2^+_2, 3^+_1$ and 4^+_2) and the two neutron separation energies of all isotopes in each isotopic chain. Typically, η decreasing from 1 to 0 as boson number increases and the nuclei evolve from vibrational to rotational as expected. This trend is observed for the studied isotopic chains and illustrated in figures (4-7) by plotting the PES from Hamiltonian (12) as a function of quadruple deformation parameter β for different values of the

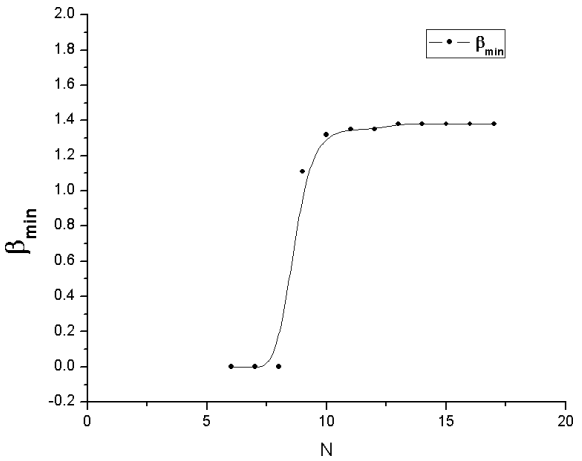

Fig. 8: Position of the absolute minima β_{min} versus the total number of bosons N from N = 6 to N = 17.

Table 1: Neutron Number.

Nucleus	η/N_{crit}			
$_{66}$Dy	0.08183	0.07339	0.04166	0.00993
$_{64}$Gd	0.08183	0.07339	0.04166	0.00993
$_{62}$Sm	0.0982	0.08807	0.5	0.01192
$_{60}$Nd	0.10911	0.09786	0.55555	0.01324
N	84	86	88	90

Nucleus	η/N_{crit}			
$_{66}$Dy	0.00149	0.0002	0.0000	0.0000
$_{64}$Gd	0.00149	0.0002	0.0000	0.0000
$_{62}$Sm	0.00179	0.00024	0.00003	0.0000
$_{60}$Nd	0.00199	0.00027		
N	92	94	96	98

the control parameter η and varying boson number N.

Here, we observe that the transition from spherical to prolate deformed occurs between N=9 and N=12. In the $^{144-154}$Nd, the nuclei $^{146-150}$Nd are transitional isotopes between the spherical nucleus ^{144}Nd and the well prolate deformed nuclei $^{152-154}$Nd. The ^{150}Gd nucleus still shows a vibrational structure while $^{156-162}$Gd are considered as rather good SU(3) example.

The $^{158-162}$Gd are corresponds to $\eta = 0$. One can observe a sudden transition in the Gd isotopes from a vibrational region into the rotational SU(3) limit. The control parameter η for each nucleus is shown in Table (1). The position of the absolute minimum $\beta_{min}(N)$ of the different PES's is illustrated in Figure (8).

Table (1) lists values of the control parameter η/N_{crit} for each Nd/Sm/Gd/Dy isotopic chain as a function of the neutron number.

5 Conclusion

In the present paper we have analyzed systematically the *PES's* for the even-even Nd/Sm/Gd/Dy isotopes using the simplified form of IBM in its sd-boson interaction. We have analyzed the critical points of the shape phase transitional region U(5)-SU(3) in the space of two control parameters η and χ.

In all isotopic chains one observes a change from spherical U(5) shape to axially symmetric deformed shape SU(3) when moving from the lighter to the heavier isotopes.

Submitted on April 8, 2013 / Accepted on April 19, 2013

References

1. Iachello F. and Arima A. The Interacting Boson Model. Cambridge University Press, Cambridge, England, 1987.
2. Frank A. and VanIsacker P. Algebraic Methods in Molecular and Nuclear Structure Physics. Wiley, New York, 1994.
3. Arima A. and Iachello F. Interacting Boson Model of Collective States I: The Vibrational Limit. *Annals of Physics*, 1976, v. 99, 253–317.
4. Arima A. and Iachello F. Interacting Boson Model of Collective States II: The Rotational Limit. *Annals of Physics*, 1978, v. 101, 201–238.
5. Arima A. and Iachello F. Interacting Boson Model of Collective States IV: The O(6) limit. *Annals of Physics*, 1979, v. 102, 123–468.
6. Casten R.E. Nuclear structure from a simple persective. Oxford University, Oxford, 1990.
7. Jolie J. et al. Quantum Phase Transition for γ-Soft Nuclei. *Physical Review Letters*, 2001, v. 87, 162501–162504.
8. Jolie J. et al. Triple Point of Nuclear Deformation. *Physical Review Letters*, 2002, v. 89, 182502–182504.
9. Jolie J. et al. Shape phase transitions in odd-mass nuclei using a supersymmetric approach. *Physical Review*, 2002, v. C70, 011305R–011308R.
10. Jolie J. et al. Two-Level Interacting Boson Models Beyond The Mean Field. *Physical Reveiw*, 2007, v. C75, 014301R - 014310R.
11. Iachello F. and Zamfir N.V. Quantum Phase Transitions in Mesoscopic Systems. *Physical Review Letters*, 2004, v. 92(3), 212501–212504.
12. Cejnar P. Shape phase transitions in rotating nuclei via cranking the interacting boson model. *Physical Review*, 2002 v. C65, 044312–044320.
13. Cejnar P., Heinze S., and Jolie J. Ground-State Shape Phase Transitions in Nuclei: Thermodynamic analogy and fitie N-effects. *Physical Review*, 2003 v. C68, 034326-034326.
14. Cejnar P. Landau Theory of Shape Phase Transitions in the Cranked Interacting Boson Model. *Physical Review Letters*, 2003, v. 90, 112501–112504.
15. Cejnar P., Heinze S. and Dobes J. Thermodynamic Analogy for Quantum Phase Transitions at Zero Temperature. *Physical Review*, 2005, v. C71, 011304R–011309R.
16. Rowe D.J., Turner P.S. and Rosensteel G. Scaling Properties and Asymptotic Spectra of Finite Models of Phase Transitions as They Approach Macroscopic Limits. *Physical Review Letters*, 2004, v. 93, 232502–232505.
17. Rowe D.J. Quasi Dynamical Symmetry in an Interacting Boson Model Phase Transition. *Physical Review Letters*, 2004, v. 93, 122502–122505.
18. Liu Y.X., Mu L.Z. and Wei H. Approach to The Rotation Driven Vibrational to Axially Rotational Shape Phase Transition Along The Yrast Line of a Nucleus. *Physics Letters*, 2006, v. B633, 49–53.

19. Zhang Y., Hau Z. and Liu Y.X. Distinguishing a First Order From a Second Order Nuclear Shape Phase Transition in The Interacting Boson Model. *Physical Review*, 2007, v. C76, 011305R–011308R.
20. Arios J.M., Dukelsky J. and Garcia-Ramos J.E. Quantum Phase Transitions in the Interacting Boson Model: Integrability, Level Repulsion and Level Crossing. *Physical Review Letters*, 2003, v. 91, 162502–162504.
21. Garcia-Ramos J.E. et al. Two-Neutron Separation Energies, Binding Energies and Phase Transitions in The Interacting Boson Model. *Nuclear Physics*, 2001, v. A688, 735–754.
22. Liu M.L. Nuclear Shape-Phase Diagrams. *Physical Review*, 2007, v. C76, 054304–054307.
23. Heyde K. et al. Phase Transitions Versus Shape Coexistence. *Physical Review*, 2004, v. C69, 054304–054309.
24. Iachello F. Dynamic Symmetrcies at The Critical Point. *Physical Review Letters*, 2000, v. 85, 3580–3583.
25. Iachello F. Analytic Prescription of Critical Point Nuclei in a Spherical Axially Deformed Shape Phase Transtion. *Physical Review Letters*, 2001, v. 87, 052502–052506.
26. Iachello F. Phase Transition an Angle Variables. *Physical Review Letters*, 2003, v. 91, 132502-132507.
27. Caprio M.A. and Iachello F. Analytic descriptions for transitional nuclei near the critical point. *Nuclear Physical*, 2007, v. A781, 26–66.
28. Arios J.M. E2 transitions and quadrupole moments in the E(5) symmetry. *Physical Review* , 2001, v. 63, 034308–034312.
29. Caprio M.A. Finite well solution for the E(5) Hamiltonian. *Physical Review*, 2002, v. C65, 031304R–031307R.
30. Caprio M.A. Consequences of wall stiffness for a β-soft potential. *Physical Review*, 2004, v. C69, 044307 - 044314.
31. Fortunato L. and Vitturi A. New analytic solutions of the collective Bohr Hamiltonian for a β-soft, γ-soft axial rotor. *Journal of Physics G: Nuclear and Particle Physics*, 2004, v. 30, 627–632.
32. Fortunato L. Soft triaxial rotovibrational motion in the vicinity of $\gamma = \pi/6$. *Physical Review*, 2004, v. 70, 011302R–011305R.
33. Zamfir N.V. et al. ^{102}Pd: An E(5) nucleus? *Physical Review*, 2002, v. C51, 044325–044332.
34. Clark R.M. et al. Searching for E(5) behavior in nuclei. *Physical Review*, 2004, v. C69, 064322–064328.
35. Krücken R. et al. B(E2) Values in ^{150}Nd and the Critical Point Symmetry X(5). *Physical Reviw Letters*, 2003, v. 88, 232501 - 232504.
36. Clark R.M. et al. Searching For X(5) Behavior in Nuclei. *Physical Review*, 2003, v. C68, 037301–037304.
37. Tonev D. et al. Transition Probabilities in ^{154}Gd: Evidence for X(5) Critical Point Symmetry. *Physical Review*, 2004, v. C69, 034334–034340.
38. McCutchan E.A. et al. Low Spin States in ^{162}Yb and The X(5) Critical Point Symmetry. *Physical Review*, 2004, v. C69, 024308–024317.
39. Dewald A. et al. Test of the critical point symmetry X(5) in the mass A=180 region. *Journal of Physics G: Nuclear and Particle Physics*, 2005, v. 31, S1427.
40. Pietralla N. and Gorbachentso O.M. Evolution of the "β excitation" in axially symmetric transitional nuclei. *Physical Review*, 2004, v. C70 011304R–011308R.
41. Bonatsos D. et al. X(3): an exactly separable γ-rigid version of the X(5) critical point symmetry. *Physical Letters*, 2006, v. B632, 238–242.
42. Bonatsos D. et al. γ-rigid solution of the Bohr Hamiltonian for $\gamma = 30^\circ$ compared to the E(5) critical point symmetry. *Physical Letters*, 2005, v. B621, 102–108.
43. Bonatsos D. et al. Z(5): critical point symmetry for the prolate to oblate nuclear shape phase transition. *Physical Letters*, 2004, v. B588, 172–179.
44. Levoi G. and Arias J. M. The sextic oscillator as a γ-independent potential. *Physical Review*, 2004, v. C69, 014304-014309.
45. Bohr A. and Mottelson. Nuclear Structure v. II. Benjamin, New York, 1975.
46. Ginocchio J. N. and Kirson M.W. Relationship between the Bohr Collective Hamiltonian and the Interacting-Boson Model. *Physical Reviw Letters*, 1980, v. 44, 1744–1747.
47. Dieperink A.E.L., Scholten O. and Iachello F. Classical Limit of the Interacting Boson Model. *Physical Reviw Letters*, 1980, v. 44, 1747–1750.

The Self-Gravity Model of the Longitudinal Span of the Neptune Arc Fraternité

K.H. Tsui, J.A. Souza, and C.E. Navia
Instituto de Fisica - Universidade Federal Fluminense, Campus da Praia Vermelha,
Av. General Milton Tavares de Souza s/n, Gragoata, 24.210-346, Niteroi, Rio de Janeiro, Brasil.
E-mail: tsui@if.uff.br

According to recent work [13, 14], the Neptune Adams ring main arc Fraternité is regarded as captured by the corotation elliptic resonance (CER) potential of Galatea. The minor arcs Egalité (2,1), Liberté, and Courage are located at positions where the time averaged forces, due to the 42-43 corotation-Lindblad resonances under the central field of Neptune, vanish. With adequately chosen Fraternité mass and Galatea eccentricity, this model gives minor arc locations compatible to observed positions, and allows a dynamic transport of materials among arcs. To complement this model, the effect of self-gravity of Fraternité, with a distributed mass, is evaluated together with the CER potential to account for its 10^o longitudinal span. Although self-gravity is the collective action of all the particles in the arc, each individual particle will see the self-potential with a central maximum as an external potential generated by other particles.

1 Introduction

From the very first observations of the Neptune Adams ring arcs [6, 12], plus the subsequent observations [2, 11], the Adams arcs seemed to change in arc locations and in brightness. More recently, these dynamic natures of the arcs, Fraternité, Egalité (2,1), Liberté, and Courage, have been confirmed beyond any doubt in another ground observation [1]. Measuring from the center of the main arc Fraternité, they extend a total of about 40^o ahead of Fraternité. Occasionally, some arcs flare up and others fade away. Furthermore, the arc configuration appears to be changing in time as well. The leading arc Courage appears to have leaped over to another CER site recently [1]. Although the twin arc Egalité (2,1) is small, it is a very bright arc. According to de Pater et al [1], its relative intensity to Fraternité varied from 17 percent higher in 2002 to seven percent lower in 2003 totaling a 24 percent relative change over a short period of time. The angular span of the twin arc Egalité appeared to be 30 percent larger in 2005 and 1999 publications than in 1989 Voyager 2 results. This widening of Egalité was accompanied by a corresponding narrowing of Fraternité, which indicated a likely exchange of material between the two. As for Liberté, 1999 data showed it was about 3^o ahead of its position in Voyager 2 pictures. For the 2005 results, the 2002 data appeared to show Liberté as a twin arc separated by about 4.5^o with the leading twin at the original Voyager 1989 location, while in 2003 it returned again as one single arc at the Voyager location. With respect to the normally low intensity arc Courage, it flared in intensity to become as bright as Liberté in 1998 indicating a possible exchange of material between the two arcs. Most interestingly, it was observed in the 2005 data that Courage has moved 8^o ahead from 31.2^o to 39.7^o [1].

According to the prevailing theories, based on the restricted three-body framework (Neptune-Galatea-arcs) with a conservative disturbing potential, these arcs are radially and longitudinally confined by the corotation resonance potential of the inner moon Galatea. In order to account for these arcs, the 84/86 corotation resonance due to the inclination of Galatea (CIR) had been invoked to give a potential site of 4.18^o [4]. Later on, because of its eccentricity (CER), the 42/43 resonance was considered giving a resonant site of 8.37^o on the Adams ring arcs [3, 5, 10]. The arc particles librate about the potential maximum imposed by the corotational resonance satellite Galatea. Dissipated energy of the particle is replenished by the Lindblad resonance. Nevertheless, well established as it is, there are several difficulties. Firstly, with Fraternité centered at the potential maximum spanning approximately 5^o on each side, it crosses two unstable potential points which ought to reduce the angular spread. Secondly, the minor arcs leading ahead of Fraternité are mislocated with the CIR or CER potential maxima. Furthermore, should the arcs were confined by the corotation potential, there ought to be arcs in other locations along the Adams ring distributed randomly instead of clustered near Fraternité.

2 Time-dependent arcs

Recently, there is a model that considers Fraternité as being captured by the CER potential of Galatea. With Fraternité having a finite mass, the minor arcs are clustered at locations along the Adams ring where the time averaged force vanishes under the corotation-Lindblad resonances [13, 14]. The finite mass of Fraternité has been suggested by Namouni [9] and Porco [10] to pull on the pericenter precession of Galatea to account for the mismatch between the CER pattern speed and the mean motion of the arcs. The arc locations are determined by the Lindblad resonance reaction of the arc itself. Because the force vanishes only on a time averaged base, as comparing to the stationary CER potential in the rotating frame, the arc material could migrate on a long time scale from one site to another leading to flaring of some arcs and fading of oth-

ers. This could also generate twin arcs (Egalité, Liberté) and displace Courage from 31.2° to 39.7° (resonant jump) [1], as required by observations. Although there are only arcs in the leading positions ahead, arcs in the trailing positions behind could be allowed in this model. According to this Lindblad reaction model, only Fraternité f is confined by the externally imposed CER potential of Galatea x which reads

$$\Phi_c \;=\; \frac{Gm_x}{a_x}\,\frac{1}{2}\left(2n + a_x\frac{\partial}{\partial a_x}\right)\frac{1}{a_x}\,b^{(n)}_{1/2}(\alpha)\,e_x\cos\phi_{fx}, \quad (1)$$

where $\vec{r}_x = (r_x, \theta_x)$ and $\vec{r} = (r, \theta)$ are the position vectors of Galatea x and Fraternité mass distribution, a_x and a are the respective semi-major axes, ϕ_x and e_x are the arguments of perihelion and eccentricity of x, $\phi_{fx} = (n\theta - (n-1)\theta_x - \phi_x)$ is the corotation resonance variable, $b^{(n)}_{1/2}(\alpha)$ is the Laplace coefficient, $\alpha = a_x/a < 1$, and $n = 43$. With $a_x = 61952.60$ km, $a = 62932.85$ km, and $\alpha = 0.98444$ [2, 11], the CER potential is

$$\Phi_c \;=\; \frac{Gm_x}{a_x}\,34\,e_x\cos\phi_{fx}. \quad (2)$$

To complement this model, we consider the self-gravity of Fraternité, which has a distributed mass, on the CER potential to account for its longitudinal 10° arc span. We first consider a qualitative spherical self-gravity physical model to grasp the 10° arc span. We begin with the Gauss law of the gravitational field

$$\nabla\cdot\vec{g}(\vec{r}) \;=\; -4\pi G\rho(\vec{r}), \quad (3)$$

$$\vec{g} \;=\; +\nabla\Phi. \quad (4)$$

Under a qualitative physical model of arc span, we take a spherical uniform mass distribution of radius r_0. Solving for the potential $\Phi(r_*)$ inside the sphere with $\rho(\vec{r}) = \rho_0$ and outside the sphere with $\rho(\vec{r}) = 0$ respectively, where r_* is measured from the center of Fraternité, and matching the potential and the gravitational field across the boundary, we get

$$\Phi_f \;=\; -\frac{1}{2}\frac{Gm_f}{r_0}\left(\frac{r_*}{r_0}\right)^2 + \frac{3}{2}\frac{Gm_f}{r_0}, \quad 0 < r_* < r_0, \quad (5)$$

$$\Phi_f \;=\; +\frac{Gm_f}{r_*}, \quad r_0 < r_* < \infty. \quad (6)$$

This potential shows a normal $1/r_*$ decaying form for $r_0 < r_*$, but a r_*^2 form for $r_* < r_0$. Writing in terms of a_x and m_x, we have for $0 < r_* < r_0,\ \delta\theta < \delta\theta_0$,

$$\begin{aligned}\Phi_f \;&=\; -\frac{1}{2}\,\frac{Gm_f}{a_x}\,\frac{a_x}{r_0}\left(\frac{r_*}{r_0}\right)^2 + \frac{3}{2}\,\frac{Gm_f}{r_0}\\ &=\; -\frac{1}{2}\,\frac{Gm_x}{a_x}\,\frac{m_f}{m_x}\,\frac{a_x}{r_0}\left(\frac{a}{r_0}\right)^2(\delta\theta)^2 + \frac{3}{2}\,\frac{Gm_x}{a_x}\,\frac{m_f}{m_x}\,\frac{a_x}{r_0},\end{aligned} \quad (7)$$

and for $r_0 < r_* < \infty,\ \delta\theta_0 < \delta\theta$,

$$\Phi_f \;=\; +\frac{Gm_f}{a_x}\,\frac{a_x}{r_*} \;=\; +\frac{Gm_x}{a_x}\,\frac{m_f}{m_x}\,\frac{1}{\delta\theta}, \quad (8)$$

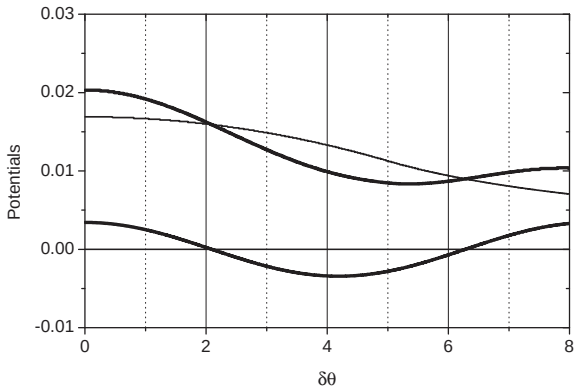

Fig. 1: The CER sinusoidal potential of Galatea in thick line, the self-potential of Fraternité with spherical model in thin line, and the sum of the two in thick line are plotted in units of Gm_x/a_x.

where r_* is now taken on the longitudinal direction along the arc, so that we can write $r_* = a\delta\theta$ and $r_0 = a\delta\theta_0$ with $\delta\theta$ as the angular span in radian. Taking $m_f/m_x = 10^{-3}$, $e_x = 10^{-4}$, and $\delta\theta_0 = 5^\circ = 0.087$ rad, which are within the estimates of the arc parameters [9], we have plotted in Fig. 1 the sinusoidal CER potential in thick line with a minimum around $\delta\theta = 4^\circ$ and the self-potential in thin line in units of Gm_x/a_x. The superposition of the two in thick line is also shown in the same figure. The superimposed potential has a maximum at the center and a minimum around $\delta\theta = 5^\circ$. Although self-gravity is resulted from all the particles of the arc, each individual particle will see the self-potential as an external potential. The particles will girate in stable orbit about the central maximum of the superpositioned CER potential and self-potential.

3 Self-gravity

We now present an elongated ellipsoid model of self-gravity. For an ellipsoidal mass distribution with uniform density ρ_0 over a volume

$$\left(\frac{x}{a_1}\right)^2 + \left(\frac{y}{a_2}\right)^2 + \left(\frac{z}{a_3}\right)^2 \;=\; 1\,, \quad (9)$$

where $a_1 > a_2 > a_3$, the potential in space for the gravitational field $\vec{g}(\vec{r})$ have been addressed in honorable treatises such as Kellogg [7] and Landau and Lifshitz [8]. Here, we follow the celebrated original work of Kellogg [7] especially in Section 6 of Chapter 7. The potential in space of this homogeneous ellipsoid is given by

$$\begin{aligned}\Phi_f(x,y,z) \;&=\; G\rho_0\pi a_1a_2a_3\times\\ &\int_0^\infty \frac{d\lambda}{\sigma^{1/2}(\lambda)}\left[1 - \frac{x^2}{a_1^2+\lambda} - \frac{y^2}{a_2^2+\lambda} - \frac{z^2}{a_3^2+\lambda}\right],\end{aligned} \quad (10)$$

where

$$\sigma(\lambda) \;=\; \left(a_1^2+\lambda\right)\left(a_2^2+\lambda\right)\left(a_3^2+\lambda\right), \quad (11)$$

and where λ parameterizes a family of ellipsoids. Consider a prolate ellipsoid with $a_1 > a_2 = a_3$. This ellipsoid has a circular cross section on the y-z plane and an axis of symmetry in x. The y-z plane of $x = 0$ is the equatorial plane. In this prolate case, the self-potential inside and outside the ellipsoid is given respectively by [7, Exercise 6, p.196]

$$\Phi_f(x,r) = G\rho_0 \frac{4\pi}{3} a_1 a_2 a_3 \frac{1}{f^2} \times \left[\left(4x^2 - 2r^2 - f^2\right) \frac{1}{2f} \ln\left(\frac{2a_1 - f}{2a_1 + f}\right)^{1/2} + \frac{4a_1^2\left(2x^2 - r^2\right) - 2f^2x^2}{2a_1\left(4a_1^2 - f^2\right)} \right], \quad (12)$$

$$\Phi_f(x,r) = G\rho_0 \frac{4\pi}{3} a_1 a_2 a_3 \frac{1}{f^2} \times \left[\left(4x^2 - 2r^2 - f^2\right) \frac{1}{2f} \ln\left(\frac{s - f}{s + f}\right)^{1/2} + \frac{s^2\left(2x^2 - r^2\right) - 2f^2x^2}{s\left(s^2 - f^2\right)} \right], \quad (13)$$

where

$$\left(\frac{f}{2}\right)^2 = a_1^2 - a_2^2,$$
$$r^2 = y^2 + z^2,$$

f is the distance between the two foci, r is the perpendicular distance to the axis of symmetry, s is the sum of distances from the two foci to the point of interest $\vec{r}$. The inside potential can be obtained from the outside potential by using $s = 2a_1$. To evaluate the potential on the axis of symmetry, we take $r = 0$. Denoting $m_f = \rho_0(4\pi/3)a_1a_2a_3$ and considering $a_1 \gg a_2$, we get

$$\Phi_f(x) = -\frac{Gm_f}{a_x}\frac{a_x}{4a_1}\left[\ln\left(\frac{2a_1}{a_2}\right) - 1\right]\left(\frac{\delta\theta}{\delta\theta_0}\right)^2 + \frac{Gm_f}{a_x}\frac{a_x}{4a_1}\ln\left(\frac{2a_1}{a_2}\right), \quad \delta\theta < \delta\theta_0, \quad (14)$$

$$\Phi_f(x) = -\frac{Gm_f}{a_x}\frac{a_x}{4a_1}\left[\frac{1}{2}\ln\left(\frac{x + f/2}{x - f/2}\right)\left(\frac{\delta\theta}{\delta\theta_0}\right)^2 - \left(\frac{\delta\theta}{\delta\theta_0}\right)\right] + \frac{Gm_f}{a_x}\frac{a_x}{4a_1}\frac{1}{2}\ln\left(\frac{x + f/2}{x - f/2}\right), \quad \delta\theta_0 < \delta\theta, \quad (15)$$

for the self-potential inside and outside the ellipsoid respectively. Taking again $m_f/m_x = 10^{-3}$, with $a_x = 61952.60$ km for Galatea, and semi-major axes $a_1 = 5500$ km and $a_2 = 55$ km, the CER potenial, the self-potential, and the superposition of the two with a minimum around $\delta\theta = 5^o$ are shown in Fig. 2.

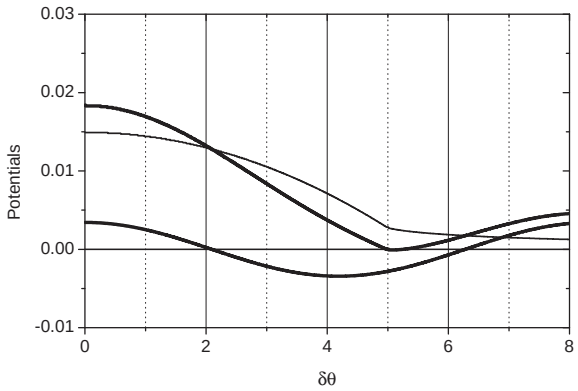

Fig. 2: The CER sinusoidal potential of Galatea in thick line, the self-potential of Fraternité with ellipsoidal model in thin line, and the sum of the two in thick line are plotted in units of Gm_x/a_x.

With Fraternité 1×10^{-3} of the mass of Galatea, the self-potential actually exceeds the CER potential in magnitude, as shown in Fig. 2. Each test mass would be librating around the potential maximum, dominated by the self-gravity of the collective mass distribution. Should Fraternité be elongated further while maintaining the total mass, it would increase the semi-major axis a_1 of the ellipsoid. This would reduce the amplitude of the self-potential of (14) through the $(a_x/4a_1)$ factor in the constant term, and weaken the self-potential. The elongation would feed the minor arcs. With this self-gravity model, not just the minor arcs are dynamically changing [1], the main arc Fraternité could be under a dynamical process as well.

4 Conclusions

In order to explain the 10^o arc span of Fraternité, we draw attention to the fact that Fraternité, as an arc, has a significant mass. This mass is a distributed mass, instead of a point-like mass, such that its self-gravity should be taken into considerations to account for its angular span. We have used two models to evaluate the self-potential in the longitudinal direction. First is the tutorial spherical model, as a proof of principle study, with a uniform mass distribution over a sphere of radius r_0. Second is the elongated ellipsoidal model for a more realistic evaluation. Using the accepted range of Fraternité parameters, the ellipsoid model shows that the self-potential of the arc could be the cause of its angular span. For a longer arc, the ellipsoid gets longer and the ratio a_1/a_2 becomes larger. Eventually, for a complete ring, the ellipsoid is infinitely long and the self-potential in the longitudinal direction becomes constant. The effects of self-gravity are felt only in the transverse direction for a planetary ring.

Submitted on March 4, 2013 / Accepted on April 9, 2013

References

1. de Pater I., et al. The dynamic neptunian ring arcs: evidence for a gradual disappearance of Liberte and resonant jump of Courage. *Icarus*, 2005, v. 174, 263–272.

2. Dumas C., et al. Stability of Neptune's ring arcs in question. *Nature*, 1999, v. 400, 733–735.
3. Foryta D. W. and Sicardy B. The dynamics of the Neptunian Adams ring's arcs. *Icarus*, 1996, v. 123, 129–167.
4. Goldreich P., Tremaine S., and Borderies N. Towards a theory for Neptune's arc rings. *Astronomical Journal*, 1986, v. 92, 490–494.
5. Horanyi M. and Porco C. C. Where exactly are the arcs of Neptune? *Icarus*, 1993, v. 106, 525–535.
6. Hubbard W. B., et al. Occultation detection of a neptunian ring like arc. *Nature*, 1986, v. 319, 636–640.
7. Kellogg O. D. Foundations of Potential Theory. Dover Publications, New York, 1953.
8. Landau L. D. and Lifshitz E. M. The Classical Theory of Fields. Pergamon Press, Oxford, 1975.
9. Namouni F. and Porco C. C. The confinement of Neptune's ring arcs by the moon Galatea. *Nature*, 2002, v. 417, 45–47.
10. Porco C. C. An explanation for Neptune's ring arcs. *Science*, 1991, v. 253, 995–1001.
11. Sicardy B., et al. Images of Neptune's ring arcs obtained by a ground based telescope. *Nature*, 1999, v. 400, 731–732.
12. Smith B. A., et al. Voyager 2 at Neptune: imaging science results. *Science*, 1989, v. 246, 1422–1449.
13. Tsui K. H. The configuration of Fraternite-Egalite2-Egalite1 in the Neptune ring arcs system. *Planetary Space Science*, 2007, v. 55, 237–242.
14. Tsui K. H. The dynamic nature of the Adams ring arcs - Fraternite, Egalite (2,1), Liberte, Courage. *Planetary Space Science*, 2007, v. 55, 2042–2044.

Probabilistic Factors as a Possible Reason of the Stability of Planetary and Electronic Orbits

Anatoly V. Belyakov
E-mail: belyakov.lih@gmail.com

An explanation is proposed that probabilistic factors cause the existence of the stable planetary orbits and electronic ones. It is confirmed when constructing frequency distributions of relevant virials.

Why there are stable planetary orbits and electronic ones, and how are they formed at all? This is all the more incomprehensible because the centrifugal forces and gravitational forces (or electrostatic ones for the atom) have a different dependence on the distance that leads only to *an unstable equilibrium*. Sure, there are some hidden factors, they may be probabilistic ones.

Thus, K. I. Dombrowski has revealed a possible connection of the sizes of planetary orbits to density of rational numbers on the number axis [1]. On the other hand, S. E. Shnoll has experimentally observed dependence of the fine structure of the normal distributions of various physical processes upon the algorithms that determine these processes [2]. It can be assumed that discrete nature of the normal distributions (and, apparently, any others) has a fundamental character.

An array of numbers that are the result of some computation algorithm can be analyzed by means of the frequency distribution*. As an example, one considers the distribution of orbits in the Bohr's atom planetary model and in the solar system.

It is known the orbital radii of the electron in the Bohr's atom to be proportional to the squares of integers. Though the existence of the orbits, i.e. the certain electronic levels, is due to quantum laws, however, this fact can also be explained by probabilistic factors.

According to the Bohr's model and proceeding from the balance of the Coulomb's and centrifugal forces, the orbital radii of the electron are in the simplest case proportional to expression $(z/\nu)^2$, where z can be regarded as a geometric mean value between the number of the elementary charges of a nucleus and electrons interacting with each other, and ν is the orbital velocity of the electron in some dimensionless units.

Let z and ν take arbitrary values, for example, from 1 to 100. Then the frequency distribution of the array of values of the function $(z/\nu)^2$ has the form shown in Fig. 1.

One can see that the peaks of the first order along the Y-axis (i.e. the most probable value) have next in values of the function $(z/\nu)^2$ along the X-axis: 1, 4, 9, 16, etc., that is, orbital radii in the Bohr's atom are proportional to the squares of integers, i.e. to the squares of electronic orbit numbers. Such distributions (or quadratic parts thereof) were also found in other, more complicated cases.

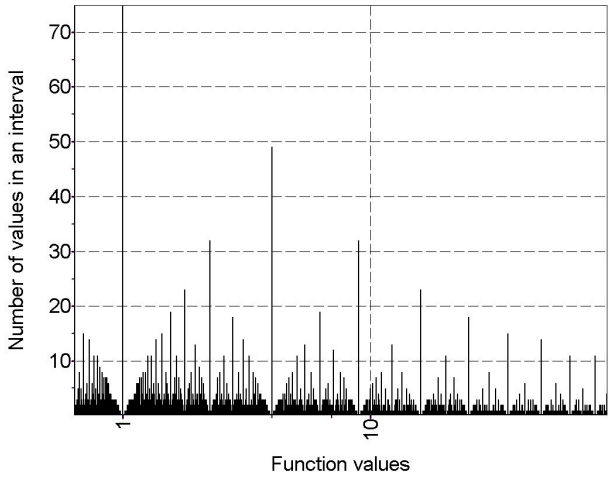

Fig. 1: Frequency distribution obtained with number of the numerical values in the scale 9,800 (of those, nonzero intervals are 3,300).

Let one consider the distribution of the planetary orbits in the solar system. Their stability can to some extent be explained by the phenomenon of orbital resonance, but this explanation is certainly not enough. As for the well-known Titius-Bode formula, then it should not be found in any of the known laws.

The equation relating the orbital radius of a planet R_0, its orbital velocity ν_0 and the mass M of a central body is:

$$R_0 = \frac{\gamma M}{\nu_0^2}, \qquad (1)$$

where γ is the gravitational constant.

In this case it would seem the frequency distribution for the orbit positions cannot be built because the function has only one variable argument ν_0, while others are permanent. However, one can assume that during formation of the solar system the mass of the central body has not been equivalent to a point with a mass equal to the mass of the Sun, and other disturbing factors could have been.

*Frequency distributions provide a possibility for bonding the probability of the appearance of numerical values of a function in the area where it exists. That is, the frequency distributions show the reproducibility of numerical values of the function due to allowed varying its arguments. There is a ready-to-use function "frequency" in Excel©; any other software can be applied as well.

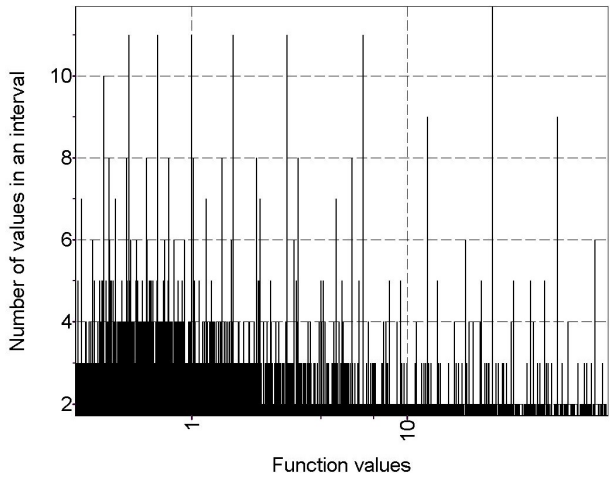

Fig. 2: Frequency distribution obtained with number of the numerical values in the scale 110,000 (of those, nonzero intervals are 48,800), $j = 0.5 \ldots 1.8$, $\nu = 0.02 \ldots 2$.

Therefore one can introduce a varied factor j in the formula and write (1) as follows:

$$R = \frac{j}{\nu^2}\,, \tag{2}$$

where R is the radius of the planetary orbit in astronomical units (a.u.), ν is the orbital velocity in the units of the orbital velocity of the Earth.

Fig. 2 shows an example of the frequency distribution of the array of values of the function (2) at $j = 0.5 \ldots 1.8$ with a step 0.025 and at $\nu = 0.05 \ldots 2$ with a step 0.01. Although the form of the distribution depends on the range of variation j and ν, the number of intervals they are divided, split range mode (step-by-step or random), and the number of processed values, but in all cases the amplitude peaks or the frequency concentrations are revealed on graphs.

In Fig. 2 from left to right the peaks of the first order (the highest) are located at the radii (in a.u.): 0.39, 0.50 (a possible orbit), 0.70, 1.0, 1.55, 2.75, 6.2, 12.3, 18.7 (a second-order peak), 25, 31 (a second-order peak), 50, and 74. Moreover, most of the values are in good agreement with the actual orbital radii of the planets. In comparison, their actual values are: 0.39, 0.72, 1, 1.52, 2.5-3.0, 5.2, 9.54, 19.2, 30.6, 30–50, 38–98, including the asteroids orbit (2.5–3.0) and the tenth planet orbit (38–98).

Of course, such simple simulation can not give a complete numerical coincidence. The more important thing is a possibility for the frequency distributions to determine the most probable values of the functions describing various processes or objects; therefore, the most stable (preferred) states of these processes or objects can also be determined [3, 4].

Submitted on April 25, 2013 / Accepted on May 10, 2013

References

1. Dombrowski K. I. Rational numbers distribution and resonance. *Progress in Physics*, 2005, v. 1, 65–67.
2. Shnoll S. E. Cosmic physical factors in random processes. Svenska fysikarkivet, Stockholm, 2009.
3. Belyakov A. V. Is the field of numbers a real physical field? On the frequent distribution and masses of the elementary particles. *Progress in Physics*, 2010, v. 3, 53–60.
4. Belyakov A. V. Finding the fine structure of the solutions of complicate logical probabilistic problems. *Progress in Physics*, 2010, v. 4, 36–39.

Double Surface and Atom Orbit

Janez Špringer

Cankarjeva cesta 2, 9250 Gornja Radgona, Slovenia, EU. E-mail: info@lekarna-springer.si

Previously (*Progr. Phys.*, 2013, v. 2, 105–106), one introduced the double surface model to explain the heterogeneous curvature of the present world. In this paper one investigates the strength of the mentioned concept in the light of forming the stable electron orbits around the atom nucleus. The conclusion is that the nature of the elliptic side of the proposed double surface offers the possibility of providing the uniform motion of the electron on the atom orbit as well as prevents the electron falling into the nucleus.

1 Theoretical background

The double surface [1] has the elliptic and hyperbolic side where the path with its translation and rotation component [2] is provided. According to this concept we have to deal with two different paths whose average is a mirror of the inverse fine structure constant. The fact that the elliptic path s can equal its translation component n [1] seems to be crucial for forming the stable electron orbits around the atom nucleus.

1.1 The elliptic side

The path on the elliptic side of the double surface can be described with the sphere law of cosines:

$$\cos\frac{s}{R} = \cos\frac{n}{R}\cos\frac{\pi}{R}\,. \tag{1}$$

On the left, s denotes the elliptic path. On the right, n and π denote the translation and rotation component of that path, respectively [2].

At $s = n$ the elliptic radius R has the potency to occupy the infinite values, since

$$\cos\frac{\pi}{R} = 1\,, \quad \text{when} \quad \frac{\pi}{R} = 2m\pi\,. \tag{2}$$

The elliptic radius expressed in Compton wavelengths of the electron is then related to the arbitrary natural number m by

$$R_{elliptic} = \frac{1}{2m}\,, \quad \text{where} \quad m \in \mathbb{N}_0\,. \tag{3}$$

For the electron only the first 431 radii are physically plausible unless one cannot imagine that the sphere could be smaller than the physical body itself. In the units of Compton wavelengths of the electron the selected elliptic radii are the next:

$$\begin{aligned} R_{elliptic} &= R_0, R_1, R_2 \cdots, R_{430} \\ &= \infty,\ \frac{1}{2},\ \frac{1}{4},\ \cdots\ \frac{1}{860} > r_{electron}\,. \end{aligned} \tag{4}$$

The greatest elliptic radius is infinite:

$$R_0 = \infty\,. \tag{5}$$

The greatest finite elliptic radius is a half of the Compton wavelength of the electron:

$$R_1 = \frac{1}{2}\,. \tag{6}$$

The smallest elliptic radius is a little bit greater than the classical electron radius itself:

$$R_{430} = \frac{1}{860} > r_{electron} = \frac{1}{2\pi\alpha^{-1}} \approx \frac{1}{861,02}\,. \tag{7}$$

1.2 The hyperbolic side

The path on the hyperbolic side of the double surface can be described with the hyperbolic law of cosines:

$$\cosh\frac{s}{R} = \cosh\frac{n}{R}\cosh\frac{\pi}{R}\,. \tag{8}$$

On the left, s denotes the hyperbolic path. On the right, n and π denote the translation and rotation component of that path, respectively [2].

According to the double surface model [1] where the characteristic values for the path and its translation component on Bohr orbit are $s = 137.072031\cdots$ and $n = 137$, the hyperbolic radius R is calculated by the equation (8) as the only one and finite:

$$R_{hyperbolic} \approx 71,520117 \text{ Compton wavelengths of the electron}\,. \tag{9}$$

2 Physical consequences on the atom level

In the double surface model Bohr radius expressed in the units of Compton wavelengths of the electron is deduced from the average path on the elliptic and hyperbolic side of the orbit:

$$R_{Bohr} = \frac{\alpha^{-1}_{elliptic} + \alpha^{-1}_{hyperbolic}}{4\pi} = \frac{\alpha^{-1}}{2\pi}\,. \tag{10}$$

The difference between $\alpha^{-1}_{observed}$ and $\alpha^{-1}_{measured}$ on the fifth decimal which was important for predicting the exact inverse fine structure previously [1], is not significant enough to be taken into account in the calculations made in this paper. From the relation (3) and (4) is seen that the radius of the elliptic side of the double surface is greater than Bohr radius only once, i.e. when $R_{elliptic} = \infty$. The infinite elliptic radius allows the

electron to move uniformly on Bohr orbit. On the other hand the 430 finite elliptic radii do not permit the electron to fall into the nucleus, because they are always much smaller than Bohr radius:

$$R_{1,2,\cdots 430} \ll R_{Bohr},$$
$$\text{since } \frac{1}{2}, \frac{1}{4}, \cdots, \frac{1}{860} < R_{Bohr} \approx 22,81. \tag{11}$$

The conclusion would be the same, if the number of the finite elliptic radii is infinite.

Thus according to the present concept the electron is closed on the elliptic sphere of the multi-sizable radius. Its destiny is to be in some way glued on Bohr orbit in the Hydrogen atom. In other atoms the similar phenomenon is expected, because their atomic radii are greater than the Bohr one and therefore still greater than the finite elliptic ones:

$$R_{atom} \geqslant R_{Bohr} \gg R_{1,2,\cdots 430}\,. \tag{12}$$

3 Conclusion

The infinite elliptic radius of the double surface enables the uniform motion of the electron on the atom orbit. The finite radii prevent the electron falling into the nucleus. From this point of view the concept of the double surface with its elliptic side as a sphere of the multi-sizable radius satisfies the demand for forming the stable electron orbits around the atom nucleus.

Respecting Plato the correct theory is only one amongst many ones revealed in the realm of the reasonable ideas.
(The Author)

Acknowledgment

Gratitude to the editors for any kind of support in lifting the veil of the physical truth.

Dedication

This fragment is dedicated to my sister Darinka.

Submitted on April 22, 2013 / Accepted on May 11, 2013

References

1. Špringer J. Double Surface and Fine Structure. *Progress in Physics*, 2013, v. 2, 105–106.
2. Špringer J. Fine Structure Constant as a Mirror of Sphere Geometry. *Progress in Physics*, 2013, v. 1, 12–14.

Multi-planet Exosystems All Obey Orbital Angular Momentum Quantization per Unit Mass predicted by Quantum Celestial Mechanics (QCM)

Franklin Potter

Sciencegems.com, 8642 Marvale Drive, Huntington Beach, CA 92646 USA. E-mail: frank11hb@yahoo.com

Quantum celestial mechanics (QCM) predicts that all orbiting bodies in gravitationally bound systems exhibit the quantization of orbital angular momentum *per unit mass*. I show that the 15 known multi-planet systems with four or more planets obey this QCM prediction. This angular momentum constraint could be the explanation for their orbital stability for billions of years, suggesting that viable models of the formation and evolution of gravitational systems must include QCM.

1 Introduction

According to recent calculations, our Solar System is unstable [1] and should have existed for only a few 100 million years! However, the Solar System has existed for more than 4.5 billion years. Obviously, some fundamental physics concept is missing. H. G. Preston and I have proposed [2] that the missing constraint is the quantization of orbital angular momentum *per unit mass* for all orbiting bodies in gravitationally bound systems. Herein I establish that all 15 known multi-planetary systems with four or more planets exhibit this constraint.

In several previous papers [2–4] we derived Quantum Celestial Mechanics (QCM) from the general theory of relativity and successfully applied QCM to numerous gravitationally bound systems, including the planets of the Solar System, the moons of the Jovian Planets, the five moons of Pluto, the Galaxy rotation velocity, gravitational lensing, clusters of galaxies, the cosmological redshift of the Universe, the circumbinary planet Kepler-16, and the S-stars at our Galaxy center.

QCM predicts that a body of mass μ orbiting a central massive object in a gravitationally bound system obeys the angular momentum L per unit mass quantization condition

$$\frac{L}{\mu} = mcH \tag{1}$$

with m an integer and c the speed of light. The Preston gravitational distance H requires only two physical parameters to determine all the possible QCM states in the system, the system's total angular momentum L_T and its total mass M_T:

$$H = \frac{L_T}{M_T c}. \tag{2}$$

In order to use this restriction, one assumes that the orbiting body is at or near its QCM equilibrium orbital radius r and that the orbital eccentricity ϵ is low so that our nearly circular orbit approximation leading to these particular equations holds true. Therefore, the L of the orbiting body will agree with its Newtonian value $L = \mu\sqrt{GM_T r(1-\epsilon^2)}$.

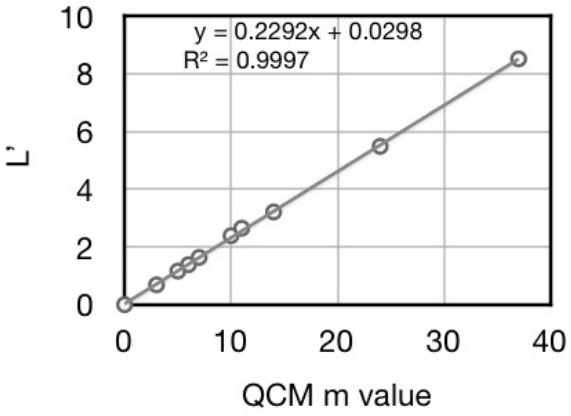

Fig. 1: The HD10180 System fit to QCM.

Every Newtonian orbit is an equilibrium orbit, but not so for QCM orbits. For a body not at the QCM equilibrium orbital radius for the QCM state or for particles near the QCM equilibrium orbital radius that could collect into a massive body, there exists a small QCM acceleration. Usually a time frame of tens or hundreds of millions of years are needed to achieve dynamic QCM equilibrium with its extremely small remnant radial oscillations. Therefore, QCM is expected to play an important role in the formation and eventual stability of multi-planetary systems over billions of years.

For circular orbits or nearly circular orbits there is a principal number $n = m + 1$ associated with the energy *per unit mass* quantization for a QCM state

$$\frac{E_n}{\mu} = -\frac{r_g^2 c^2}{8n^2 H^2} = -\frac{G^2 M_T^2}{2n^2 L_T^2} \tag{3}$$

with r_g the Schwarzschild radius of the system. The derived Schrodinger-like gravitational wave equation dictates all the physics via solutions that are hydrogen-like wave functions.

The QCM fit to the orbital parameters of all known planets of a multi-planet system determines the total angular momentum of that system, a value which can be used to predict whether more planets can be expected and/or whether

the equivalent of an Oort cloud is required. Recall that for our own Solar System the Oort Cloud dominates the total angular momentum, being a factor of at least 50 greater than the total planetary orbital momentum. Without the Oort Cloud angular momentum, QCM predicts that all the planetary orbital radii would be within the radius of the Sun! By including the angular momentum of the Oort Cloud, QCM suggests that the planets formed near to their present orbital radii.

Many exoplanetary systems have their Jupiter size planets at extremely small orbital radii, within about 1.5 AU from the star, with many more smaller planets even closer. There is the question of why such massive planets are so close to their star. One possible answer is that the system total angular momentum value is low compared to the QCM value needed to "push" the system further out. That is, QCM predicts that a larger total angular momentum for the system means larger QCM orbital spacings.

2 Multi-planetary results

Multi-planet systems are in the database called the Exoplanets Data Explorer [5], but complete data sets for HD 10180 [6], HD 40307 [7], Tau Ceti [8], GJ 676A [9], and Upsilon Andromedae [10] are only in research articles. There are hundreds of two and three planet systems which I choose to exclude herein even though they also exhibit the QCM constraint. As more planets orbiting these systems are identified, their fits to the QCM prediction can be determined.

In Table 1 are listed the host star, the star mass in solar units, the number of planets N, their QCM m values, and the slope b for $L/\mu = bx + a$ in the plot of $L' = L/\mu$ versus m for all the planets of the particular system. The plot for HD10180 is shown as an example, with the uncertainty bars for L' within the circle data points. By using both the semi-major axis and the orbital period as constraints, one obtains a linear regression fit with $R^2 > 0.999$. The system's predicted total angular momentum $L_T = b\,M_T$ multiplied by 10^{15} kg m^2/s.

From the QCM predicted L_T values, one learns that these 15 multi-planet systems have more angular momentum which is to be contributed by additional orbiting bodies such as planets and/or the equivalent of the Oort Cloud.

Table 1: QCM Multi-Planet m Values

System	Mass	N	m Values	b
HD 10180	1.055	9	3,5,6,7,10,11,14,24,37	0.23
Sun	1	8	3,4,5,6,12,17,25,31	0.77
HD 40307	0.752	6	8,11,14,17,20,29	0.10
Kepler-11	0.954	6	10,11,13,15,17,23	0.13
Kepler-20	0.912	5	14,17,21,25,40	0.06
Kepler-33	1.291	5	8,11,13,15,16	0.16
Kepler-62	0.690	5	6,8,9,17,22	0.14
Tau Ceti	0.783	5	8,11,16,19,30	0.15
55 Cancri	1.026	5	4,10,16,27,76	0.13
GJ 581	0.311	4	7,9,12,21	0.05
GJ 676A	0.71	4	3,6,20,34	0.25
GJ 876	0.334	4	5,13,17,22	0.07
HR 8799	1.472	4	12,16,20,27	1.65
Mu Arae	1.077	4	4,12,15,29	0.38
Ups Andr	1.01	4	2,10,18,27	0.42

3 Conclusions

All the 15 analyzed multi-planet systems obey the QCM orbital angular momentum per unit mass quantization condition. The integers for the m values are not sequential, implying that the history of each system plays an important role in which orbital states are occupied. For example, mass depletion in a region caused by the faster formation of a large planet might not leave enough mass for another planet to form at a nearby QCM equilibrium orbital radius.

The resulting fits are evidence that the quantization of orbital angular momentum per unit mass is an important physical factor in planetary systems and should not be ignored in studies of their formation, stability, and evolution.

Acknowledgements

The author thanks Sciencegems.com for generous support. This research has made use of the Exoplanet Orbit Database and the Exoplanet Data Explorer at exoplanets.org.

Submitted on May 18, 2013 / Accepted on May 22, 2013

References

1. Hayes W. B. Is the Outer Solar System Chaotic? arXiv: astro-ph/0702179v1.
2. Potter F. Pluto Moons exhibit Orbital Angular Momentum Quantization per Mass. *Progress in Physics*, 2012, v. 4, 3–4.
3. Preston H. G., Potter F. Exploring Large-scale Gravitational Quantization without h-bar in Planetary Systems, Galaxies, and the Universe. arXiv: gr-qc/030311v1.
4. Potter F., Preston H. G. Cosmological Redshift Interpreted as Gravitational Redshift. *Progress in Physics*, 2007, v. 2, 31–33.
5. Wright J. T. et al. The Exoplanet Orbit Database. http://adsabs.harvard.edu/abs/2011PASP..123..412W
6. Tuomi M. Evidence for 9 planets in the HD 10180 system. arXiv: 1204.1254.
7. Tuomi M., et al. Habitable-zone super-Earth candidate in a six-planet system around the K2.5V star HD 40307. arXiv: 1211.1617.
8. Tuomi M., et al. Signals embedded in the radial velocity noise. Periodic variations in the τ Ceti velocities. arXiv: 1212.4277.
9. Anglada-Escudé G., Tuomi M. A planetary system with gas giants and super-Earths around the nearby M dwarf GJ 676A. Optimizing data analysis techniques for the detection of multi-planetary systems. arXiv: 1206.7118v2.
10. Curiel S., et al. A fourth planet orbiting υ Andromedae. *Astronomy and Astrophysics*, 2011, v. 525, A78.

The Gravitational Field: A New Approach

Patrick Marquet

patrick.marquet6@wanadoo.fr

In this paper, we consider the Einstein field equations with the cosmological term. If we assume that this term is slightly varying, it induces a vacuum background field filling the space. In this case, inspection shows that the gravitational field is no longer represented by a pseudo-tensor, but appears on the right hand side of the field equations as a true tensor together with the bare mass tensor thus restoring the same conservation condition as obeyed by the Einstein tensor.

Introduction

Soon after his theory of General Relativity was published in 1916, Einstein rapidly turned to the unifying of the gravitational field with electromagnetism (which at that time was considered as the second fundamental field).

The quest for such an universal scheme ended in 1955 with the Einstein-Schrödinger theory (see for example [1]) definitely abandoned since as the quantum field theories gained the increasing successes and have been long substantiated by numerous experimental confirmations.

Basically, the unified principle adopted by the successive authors (Kaluza-Klein, Weyl, Eddington, et al.) relied either on extra dimensions, or on an extension of the Riemannian theory with additional space-time curvatures introduced to yield the electromagnetic field characteristics, and where the stress-energy tensor regarded as provisional, will be eventually absent [2, 3, 4].

Total geometrization of matter and electromagnetism was anyhow the original focus.

To understand this long period of research, one should remember that Einstein always claimed that the energy-momentum tensor (s) which can appear in the right hand side of his field equations, was "clumsy"; in short, he considered this form as an unsatisfactory solution which had to fit differently in his equations.

Einstein's argument is actually strongly supported by the following fact: while his tensor exhibits a *conceptually* conserved property, any corresponding stress energy-tensor *does not*, which leaves the theory with a major inconsistency.

When pure matter is the source, the problem has been "cured" by introducing the so-called "pseudo-tensor" that "conveniently" describes the gravitational field of this mass so that the four-momentum of both matter and its gravity field is conserved.

Unfortunately by essence this pseudo-tensor cannot appear in the field equations, and so the obvious physical defect emphasized by Einstein, still remains to-day as a stumbling block.

In this paper, we tackle this problems by proceeding as follows: in contrast to the previous theories, the energy-momentum tensor of the source is here strengthened, although we restrict our study to neutral massive flow.

In this respect, it is shown that the gravitational field of a massive body is no longer described by a *pseudo-tensor*, but appears as a *true tensor* in the field equations as it should be, in order to balance the conceptually conserved property of the Einstein tensor.

To achieve this goal we do:

- We first formulate the field equations with a massive source in density notation;
- We write the conservation law for the Einstein tensor density derived from the Bianchi identities, which cannot apply to the energy-momentum tensor density as a source;
- We then include a variable term that supersedes the so-called cosmological term Λg_{ab} in the field equations, still complying with the conservation property of the Einstein tensor density in GR;
- Under this latter assumption, we will then formally show that the gravity field of a massive source is no longer described by a vanishing *pseudo tensor* but it reduces to a true tensor describing a *persistent* vacuum background field resulting from the existence of the variable term.

1 The field equations in General Relativity

1.1 The tensor representation

In the General Theory of Relativity (GR), it is well known that by varying the action

$$\mathcal{S} = \boldsymbol{L}_E \, d^4x \,,$$

where the *Lagrangian density* is given by

$$\boldsymbol{L}_E = \sqrt{-g}\, G^{ab} \left(\begin{Bmatrix} e \\ ab \end{Bmatrix} \begin{Bmatrix} d \\ de \end{Bmatrix} + \begin{Bmatrix} d \\ ae \end{Bmatrix} \begin{Bmatrix} e \\ bd \end{Bmatrix} \right), \tag{1.1}$$

$$g = \det \| g_{ab} \| \tag{1.2}$$

one infers the *symmetric Einstein tensor*

$$G_{ab} = R_{ab} - \frac{1}{2}\, g_{ab} R, \tag{1.3}$$

where

$$R_{bc} = \partial_a \left\{ {a \atop bc} \right\} - \partial_c \left\{ {a \atop ba} \right\} + \left\{ {d \atop bc} \right\} \left\{ {a \atop da} \right\} - \left\{ {d \atop ba} \right\} \left\{ {a \atop dc} \right\} \tag{1.4}$$

is the *Ricci tensor* with its contraction R, the *curvature scalar*, while $\left\{ {e \atop ab} \right\}$ denote the Christoffel Symbols of the second kind.

The 10 *source free field equations* are

$$G_{ab} = 0. \tag{1.5}$$

The second rank Einstein tensor G_{ab} is symmetric and is only function of the metric tensor components g_{ab} and their first and second order derivatives.

The relation

$$\nabla_a G^a_b = 0 \tag{1.6}$$

is the conservation identities provided that the tensor G_{ab} has the form [5]

$$G_{ab} = k \left[R_{ab} - \frac{1}{2}\, g_{ab}(R - 2\Lambda) \right], \tag{1.7}$$

k is a constant, which is here taken 1, is usually named cosmological constant Λ.

When a source is present, the field equations become

$$G_{ab} = R_{ab} - \frac{1}{2}\, g_{ab} R - g_{ab}\Lambda = \varkappa\, T_{ab}, \tag{1.8}$$

where T_{ab} is the energy-momentum tensor of the source.

1.2 The tensor density representation

We first set

$$\boldsymbol{g}^{ab} = \sqrt{-g}\, g^{ab} \tag{1.9}$$

and the Einstein tensor density is

$$\boldsymbol{G}^{ab} = \sqrt{-g}\, G^{ab}, \quad \boldsymbol{G}^c_a = \sqrt{-g}\, G^c_a, \tag{1.10}$$

$$\boldsymbol{R}^{ab} = \sqrt{-g}\, R^{ab}. \tag{1.11}$$

In density notations, the field equations with the source (1.8) will read

$$\boldsymbol{G}^{ab} = \boldsymbol{R}^{ab} - \frac{1}{2}\, g^{ab}\boldsymbol{R} - g^{ab}\zeta = \varkappa\, \boldsymbol{T}^{ab}. \tag{1.12}$$

Here in place of the constant cosmological term Λ which should be here represented by $\Lambda\sqrt{-g}$, we have introduced a *scalar density* denoted as

$$\zeta = \Xi\sqrt{-g}. \tag{1.13}$$

Unlike Λ, the scalar Ξ is slightly variable and represents the *Lagrangian* characterizing a specific *vacuum background field* as will be shown below.

2 The conservation identities

2.1 Tensor version for the Einstein tensor

From the Bianchi identities applied to the Riemann tensor

$$R^{ci\cdot\cdot}_{\cdot\cdot bc\,;\,i} + R^{ci\cdot\cdot}_{\cdot\cdot ib\,;\,c} + R^{ci\cdot\cdot}_{\cdot\cdot ci\,;\,b} = 0 \tag{2.1}$$

we infer the conservation conditions which apply to the Einstein tensor without Ξ, and hereinafter denoted by

$$^{\circ}G^a_b = R^a_b - \frac{1}{2}\, g^a_b R. \tag{2.2}$$

The Einstein tensor thus satisfies intrinsically the conservation law:

$$\nabla_a\, {}^{\circ}G^a_b = 0. \tag{2.3}$$

2.2 Tensor density version for the Einstein tensor

In the same way, we start with the Einstein tensor density without the cosmological term

$$^{\circ}\boldsymbol{G}^{ab} = \boldsymbol{R}^{ab} - \frac{1}{2}\, g^{ab}\boldsymbol{R}. \tag{2.4}$$

With (2.3), let us write down

$$\nabla_a\, {}^{\circ}G^a_b = \partial_a\, {}^{\circ}G^a_b + \left\{ {a \atop ca} \right\} {}^{\circ}G^c_b - \left\{ {c \atop ba} \right\} {}^{\circ}G^a = \frac{\partial_a\, {}^{\circ}\boldsymbol{G}^a_b}{\sqrt{-g}} - \left\{ {c \atop ba} \right\} {}^{\circ}G^a_c = 0,$$

which is easily found to be

$$\frac{\partial_a\, {}^{\circ}\boldsymbol{G}^a_b}{\sqrt{-g}} - \frac{1}{2}\, {}^{\circ}G^{ea}\, \partial_b\, g_{ea} = 0 \tag{2.5}$$

using $dg_{ai} = -g_{ab}g_{ic}dg^{bc}$ and $dg^{ai} = -g^{ab}g^{ic}dg_{bc}$ the formula (2.5) can be also written as

$$\partial_a\, {}^{\circ}\boldsymbol{G}^a_b - \frac{1}{2}\, \boldsymbol{G}^{ea}\, \partial_b\, g_{ea} = 0. \tag{2.6}$$

The latter equation is the conservation condition for $^{\circ}\boldsymbol{G}^{ab}$ which is equivalent to (2.3).

2.3 Conservation of the energy-momentum tensor

2.3.1 Problem statement

Let us consider the energy-momentum tensor for neutral matter density ρ:

$$T_{ab} = \rho\, u_a\, u_b \tag{2.7}$$

as the right hand side of the field equations

$$^{\circ}G_{ab} = R_{ab} - \frac{1}{2}\, g_{ab} R = \varkappa T_{ab}. \tag{2.8}$$

The conservation condition for this tensor are written

$$\nabla_a T^a_b = \frac{1}{\sqrt{-g}}\, \partial_a \boldsymbol{T}^a_b - \frac{1}{2}\, T^{ac}\, \partial_b\, g_{ac} = 0 \tag{2.9}$$

with the tensor density

$$\boldsymbol{T}_b^a = \sqrt{-g}\, T_b^a. \tag{2.10}$$

However, across a given hypersurface dS_b, the integral

$$P^a = \int T^{ab}\, \sqrt{-g}\, dS_b \tag{2.11}$$

is conserved only when

$$\partial_a \boldsymbol{T}_b^a = 0. \tag{2.12}$$

From (2.6) inspection still shows that

$$\partial_a \boldsymbol{T}_b^a = \frac{1}{2}\, \boldsymbol{T}^{cd}\, \partial_b\, g_{cd} \tag{2.13}$$

but here, unlike the Einstein tensor ${}^\circ G_{ab}$ which is *conceptually conserved* ($\nabla_a {}^\circ G_b^a = 0$), the conditions

$$\nabla_a T_b^a = 0$$

or

$$\partial_a \boldsymbol{T}_b^a = 0$$

are thus never satisfied in a general coordinates system.

Therefore, the Einstein tensor ${}^\circ G_{ab}$ which *intrinsically* obeys a conservation condition, is related with a massive tensor $T_{ab}(\rho)$ which obviously *fails to satisfy the same requirement*:

$${}^\circ G_{ab} = \varkappa T_{ab}. \tag{2.14}$$

As a matter of fact, a correct formulation would consist of explicitly writing down the mass density with its gravity field, i.e. with a pseudo-tensor $(t_{ab})_{field}$.

As is known, the name *pseudo-tensor* is chosen since this quantity can be transformed away by a suitable choice of coordinates.

Hence, we should write

$$G_{ab} = \varkappa \left[(T_{ab})_{matter} + (t_{ab})_{field} \right]. \tag{2.15}$$

This is classically interpreted by requiring that the **total** 4-momentum vector P^a of *matter* with its *gravitational field*

$$P^a = \left[(T^{ab})_{matter} + t^{ab})_{field} \right] \sqrt{-g}\, dS_b \tag{2.16}$$

must be together conserved*

*Some authors [8] state that integrating $\nabla_k T_i^k = 0$ yields a conservation law for a vector $P^a = T^{ab} K_b$ when the metric admits a Killing vector $\boldsymbol{K}$: $P^a_{;a} = T^{ab}_{;a} K_b + T^{ab} K_{b;a}$ and since T^{ab} is symmetric, we have for the Lie derivative $K_{b;a} = \frac{1}{2} \boldsymbol{L}_K \boldsymbol{g}_{ab} = 0$, then $P^a_{;a} = 0$.

2.3.2 The gravity pseudo-tensor

In order to follow this way, Landau and Lifshitz [6] started from the unsuitable tensor equation (2.9)

$$\nabla_k T_i^k = \frac{1}{\sqrt{-g}}\, \partial_k \boldsymbol{T}_i^k - \frac{1}{2}\, T^{kl}\, \partial_i\, g_{kl} = 0.$$

They thus consider a special choice of a set of the coordinates which cancels out all first derivatives of the g_{ik} at a given 4-space-time point.

In this system, the energy-momentum tensor expression is given by

$$T^{ik} = \frac{1}{2\varkappa}\, \partial_e (-g)^{-1} \left[\partial_d\, (-g) \left(g^{ik} g^{ed} - g^{ie} g^{kd} \right) \right]. \tag{2.17}$$

As $\left\{ {i \atop ke} \right\}$ are postulated to be zero at the considered point, we may extract the factor $(-g)^{-1}$ from the derivative in the latter equation, so

$$(-g)\, T^{ik} = \partial_e \boldsymbol{H}^{ike} = \frac{1}{2\varkappa}\, \partial_e \left(\partial_d \boldsymbol{H}^{iked} \right).$$

The quantity

$$\boldsymbol{H}^{iked} = (-g) \left(g^{ik} g^{ed} - g^{ie} g^{kd} \right) \tag{2.18}$$

can be regarded as a "double tensor density" and is often referred to, as the "superpotential of Landau-Lifshitz" [7]. Now, in any other arbitrary system, generally

$$\partial_e \boldsymbol{H}^{ike} - (-g)\, T^{ik} \neq 0,$$

and so, we will have to bring a small tensor correction t_{LL}^{ik} (Landau-Lifshitz pseudo-tensor) which is accepted as representing the gravitational field of matter:

$$\partial_e \boldsymbol{H}^{ike} = (-g) \left(T^{ik} + t_{LL}^{ik} \right).$$

This equation implies the condition

$$\partial_k \left[(-g) \left(T^{ik} + t_{LL}^{ik} \right) \right] = 0, \tag{2.19}$$

which is the conservation law for the classical total four-momentum vector density of both matter and gravitational field written as

$$\boldsymbol{P}^i = \int \left[(-g) \left(T^{ik} + t_{LL}^{ik} \right) \right] dS_k, \tag{2.20}$$

(compare with (2.11)).

After a tedious calculation, the final form of the symmetric tensor t_{LL}^{ik} as a function of the g_{ik}, is found to be

$$\begin{aligned} (-g)\, t_{LL}^{ik} = \frac{1}{2\varkappa} \Big[& \boldsymbol{g}^{ik}_{,l} \boldsymbol{g}^{lm}_{,m} - \boldsymbol{g}^{il}_{,l} \boldsymbol{g}^{km}_{,m} + \frac{1}{2}\, g^{ik} g_{lm}\, \boldsymbol{g}^{ln}_{,p} \boldsymbol{g}^{pm}_{,n} - \\ & - \left(g^{il} g_{mn}\, \boldsymbol{g}^{kn}_{,p} \boldsymbol{g}^{mp}_{,l} + g^{kl} g_{mn}\, \boldsymbol{g}^{in}_{,p} \boldsymbol{g}^{mp}_{,l} \right) + g_{lm}\, g^{np} \boldsymbol{g}^{il}_{,n} \boldsymbol{g}^{km}_{,p} + \\ & + \frac{1}{8} \left(2 g^{il} g^{km} - g^{ik} g^{lm} \right) \left(2 g_{np} g_{qr} - g_{pq} g_{nr} \right) \boldsymbol{g}^{nr}_{,l} \boldsymbol{g}^{pq}_{,m} \Big]. \end{aligned} \tag{2.21}$$

Therefore, the Einstein field equations can be eventually written in the form:

$$\boldsymbol{H}^{iked}_{\cdots,kd} = 2\varkappa\,(-g)\left(T^{ie} + t^{ie}_{LL}\right). \tag{2.22}$$

Unfortunately, the quantity t^{ie}_{LL} which now appears on the right hand side of the field equations as it should be, is not a *true tensor*.

Hence, we are once more faced with a contradiction: the left hand side of the field equations for a massive source is a true tensor, while the right hand side is not, which reveals a major inconsistency within the theory.

2.4 Introduction of a background field tensor

Let us now try to remove this ambiguity.

We start by writing the global energy-momentum tensor density of the massive source splitting up bare matter and pure field:

$$\boldsymbol{T}^a_b = (\boldsymbol{T}^a_b)_{matter} + (\boldsymbol{t}^a_b)_{field}. \tag{2.23}$$

The field tensor density $(\boldsymbol{t}^a_b)_{field}$ is in turn composed of two parts: *gravity field + vacuum background field*

$$(\boldsymbol{t}^a_b)_{field} = (\boldsymbol{t}^a_b)_{gravity} + (\boldsymbol{t}^a_b)_{background\,field} \tag{2.24}$$

with

$$(\boldsymbol{t}_{ab})_{background\,field} = \frac{\varsigma}{2\varkappa}\, g_{ab} = \frac{\Xi\sqrt{-g}}{2\varkappa}\, g_{ab}\,. \tag{2.25}$$

According to the standard theory, we next re-formulate the field equations with a *bare* massive source

$$\boldsymbol{G}^{ab} = \boldsymbol{R}^{ab} - \frac{1}{2}\, g^{ab}\boldsymbol{R} - g^{ab}\varsigma = \varkappa(\boldsymbol{T}^{ab})_{matter} \tag{2.26}$$

under the form

$$\boldsymbol{G}^{ab} = \boldsymbol{R}^{ab} - \frac{1}{2}\, g^{ab}\boldsymbol{R} = \varkappa(\boldsymbol{T}^{ab})_{matter} + g^{ab}\varsigma. \tag{2.27}$$

3 Expliciting the field equations in density notation

3.1 Taking account of the Lagrangian Ξ

Reverting to (2.13), we now write for the *bare* matter tensor density

$$\partial_a(\boldsymbol{T}^a_b)_{matter} = \frac{1}{2}\,(\boldsymbol{T}^{cd})_{matter}\,\partial_b\, g_{cd}. \tag{3.1}$$

Inspection then shows that

$$R_{il}\, d\boldsymbol{g}^{il} = \sqrt{-g}\left[-R^{ie} + \frac{1}{2}\, g^{ie}R\right] dg_{ie} = \\ = -\varkappa(\boldsymbol{T}^{ie})_{matter}\, dg_{ie}. \tag{3.2}$$

Taking now into account the Lagrangian formulation for R_{il}, which is

$$R_{il} = \frac{\delta \boldsymbol{L}_E}{\partial \boldsymbol{g}^{il}} = \partial_k \frac{\partial \boldsymbol{L}_E}{\partial(\partial_k g^{il})} - \frac{\partial \boldsymbol{L}_E}{\partial \boldsymbol{g}^{il}}, \tag{3.3}$$

we obtain

$$-\varkappa(\boldsymbol{T}^{il})_{matter}\, dg_{il} = \partial_k \frac{\partial \boldsymbol{L}_E}{\partial(\partial_k \boldsymbol{g}^{il})} - \frac{\partial \boldsymbol{L}_E}{\partial \boldsymbol{g}^{il}}\, d\boldsymbol{g}^{il} = \\ = \partial_k \frac{\partial \boldsymbol{L}_E\, d\boldsymbol{g}^{il}}{\partial(\partial_k \boldsymbol{g}^{il})} - \partial \boldsymbol{L}_E,$$

that is

$$-\varkappa(\boldsymbol{T}^{il})_{matter}\,\partial_m g_{il} = \partial_k\left[\frac{\partial \boldsymbol{L}_E \partial_m(\partial \boldsymbol{g}^{il})}{\partial(\partial_k \boldsymbol{g}^{il})} - \delta^k_m \boldsymbol{L}_E\right] = \\ = 2\varkappa\,\partial_k(\boldsymbol{t}^k_m)_{field}, \tag{3.4}$$

where $(\boldsymbol{t}^k_m)_{field}$ denotes the field tensor density extracted from

$$2\varkappa(\boldsymbol{t}^k_m)_{field} = \frac{\partial \boldsymbol{L}_E \partial_m(\partial \boldsymbol{g}^{il})}{\partial(\partial_k \boldsymbol{g}^{il})} - \delta^k_m \boldsymbol{L}_E \tag{3.5}$$

so, that we have the explicit canonical form

$$(\boldsymbol{t}^k_m)_{field} = \frac{1}{2\varkappa}\left[\frac{\partial \boldsymbol{L}_E \partial_m(\partial \boldsymbol{g}^{il})}{\partial(\partial_k \boldsymbol{g}^{il})} - \delta^k_m \boldsymbol{L}_E\right] \tag{3.6}$$

and where

$$\partial_k(\boldsymbol{T}^k_i)_{matter} = \frac{1}{2}\,(\boldsymbol{T}^{ek})_{matter}\,\partial_k g_{ei} = -\partial_k(\boldsymbol{t}^k_i)_{field}.$$

that is, the required conservation relation

$$\partial_k\left[(\boldsymbol{T}^k_i)_{matter} + (\boldsymbol{t}^k_i)_{field}\right] = 0. \tag{3.7}$$

Then, re-instating the term ζ according to (2.24) and (2.25), the gravitational field tensor density now reads:

$$(\boldsymbol{t}^k_m)_{gravity} = \frac{1}{2\varkappa}\left[\frac{\partial \boldsymbol{L}_E \partial_m(\partial \boldsymbol{g}^{il})}{\partial(\partial_k \boldsymbol{g}^{il})}\right] - \delta^k_m(\boldsymbol{L}_E - \zeta). \tag{3.8}$$

The presence of the scalar density ζ characterizing the background field is here of central importance, as it means that $(\boldsymbol{t}^k_m)_{gravity}$ can never be zero in contrast to the classical theory, and as a result, it constitutes a *true tensor*. Such a gravity field never completely cancels out, but far from its matter source, it sharply decreases down to the level of the background field described by the tensor density $(\boldsymbol{t}^{ab})_{background\,field}$.

In addition, we clearly see that ζ represents the *lagrangian density* characterizing the background field, thus lending support to our initial hypothesis regarding the lagrangian Ξ.

In this picture, the vacuum is permanently filled with this homogeneous background energy field ensuring a smooth continuity with the gravitational field of a neighbouring mass.

3.2 Classical formulation

When the term Ξ is kept constant like the cosmological term Λ, the tensor density (3.8) reduces to

$$(\boldsymbol{t}^k_m)_{pseudogravity} = \frac{1}{2\varkappa}\left[\frac{\partial \boldsymbol{L}_E \partial_m(\partial \boldsymbol{g}^{il})}{\partial(\partial_k \boldsymbol{g}^{il})} - \delta^k_m \boldsymbol{L}_E\right], \tag{3.9}$$

which is just the classical *gravity pseudo-tensor density* that may now vanish in a given space-time point.

In this case, expressed with the explicit form of the Lagrangian density $\boldsymbol{L}_E$ written in (1.1), the expression (3.9) becomes:

$$(\boldsymbol{t}_m^k)_{pseudogravity}=\frac{1}{2\varkappa}\left[\left\{{k \atop i\,l}\right\}\partial_m \boldsymbol{g}^{il}-\left\{{i \atop i\,l}\right\}\partial_m \boldsymbol{g}^{lk}-\delta_m^k \boldsymbol{L}_E\right]. \quad (3.10)$$

This is the *mixed Einstein-Dirac pseudo-tensor density* [9] which is not symmetric on *k* and *m*, and is therefore not suitable for basing a definition of angular momentum on.

3.3 Field equations

The field equations with a massive source, which are

$$\boldsymbol{G}^{ab}=\boldsymbol{R}^{ab}-\frac{1}{2}\,g^{ab}\boldsymbol{R}-g^{ab}\zeta=\varkappa(\boldsymbol{T}^{ab})_{matter}, \quad (3.11)$$

may be now eventually re-written

$$^{\circ}\boldsymbol{G}^{ab}=\boldsymbol{R}^{ab}-\frac{1}{2}\,g^{ab}\boldsymbol{R}=\varkappa\left[(\boldsymbol{T}^{ab})_{matter}+(\boldsymbol{t}^{ab})_{gravity}\right] \quad (3.12)$$

with the explicit appearance of the gravity field as defined in (3.8) and which is now represented by a *true* tensor density.

Like we emphasized above, far from the mass, the "source free" field equations should always retain a non zero right hand side

$$^{\circ}\boldsymbol{G}^{ab}=\boldsymbol{R}^{ab}-\frac{1}{2}\,g^{ab}\boldsymbol{R}=\varkappa(\boldsymbol{t}^{ab})_{background\,field}, \quad (3.13)$$

which are the analogue of (1.7):

$$\boldsymbol{G}^{ab}=\boldsymbol{R}^{ab}-\frac{1}{2}\,g^{ab}\boldsymbol{R}-g^{ab}\zeta=0. \quad (3.14)$$

In this case, the conservation law applied to the right hand side of the tensor field equations is straightforward:

$$\nabla_a(t_b^a)_{background\,field}=\nabla_a\left(\frac{\Xi}{2\varkappa}\,\delta_b^a\right)=0, \quad (3.15)$$

from which readily follows

$$\partial_a(\boldsymbol{t}_b^a)_{background\,field}=\partial_a\left(\frac{\zeta}{2\varkappa}\,\delta_b^a\right)=0. \quad (3.16)$$

3.4 Physical description

We would like now to give a simple but instructive picture of the situation where a static mass is placed in the vacuum background energy field. Let us write the energy-momentum tensor for matter and its gravitational field as in (3.12):

$$T_{ab}=(\rho u_a u_b)_{matter}+(t_{ab})_{gravity}. \quad (3.17)$$

In virtue of the principle of equivalence, any *bare mass* of volume *V together with its gravitational field*, can be expressed through the time component of a 4-momentum P^a according to

$$P^0=\int\left(T_1^1+T_2^2+T_3^3-T_0^0\right)\sqrt{-g}\,dV, \quad (3.18)$$

where T_a^a are the skew components of the energy-momentum tensor (3.17), which implicitly contains the gravity field [10].

Now, we formulate (3.18) under the equivalent form:

$$P^0=P_0=\int\left(\boldsymbol{T}_1^1+\boldsymbol{T}_2^2+\boldsymbol{T}_3^3-\boldsymbol{T}_0^0\right)dV. \quad (3.19)$$

In the immediate vicinity of the mass, it is easy, to show that generalizing (3.19) leads to the 4-momentum vector that includes the right hand side of (3.12):

$$P_a=\int\left[(\boldsymbol{T}_a^b)_{matter}+(\boldsymbol{t}_a^b)_{gravity}\right]dS_b. \quad (3.20)$$

Far from the source, we have obviously

$$(P_a)_{background\,field}=\int\left[(\boldsymbol{t}_a^b)_{background\,field}\right]dS_b, \quad (3.21)$$

where $(\boldsymbol{t}_a^b)_{background\,field}$ is a true tensor density, and the conservation law applied to P^a holds for all configurations, in accordance with (3.7) and (3.16).

4 Conclusions and outlook

In this short paper, we have sketched here a possible way out of the gravitational field pseudo-tensor.

From the beginning of General Relativity, the cosmological constant Λ has played an unsavory role. Einstein included this constant in his theory, because he wanted to have a cosmological model of the Universe which he wrongly thought static.

But to-day, a cosmological term seems to be badly needed to explain some astronomical observed clues, within the basic dynamical expanding model of Robertson-Walker [11], even though its occurrence was never clearly explained.

However, there is no reason *à priori* to consider this cosmological term as constant everywhere.

In this respect, the background field hypothesis is rewarding in terms of several physical advantages:

- The ill-defined gravitational pseudo-tensor is now a true tensor, and it appears explicitly in the field equations with a massive source;
- The background persistent homogeneous energy field is then formally shown to be a consequence of the above derivation and it is actually regarded as the (sharply decreasing) continuation of any mass gravity field tensor;
- The inferred global energy-momentum tensor intrinsically satisfies the conservation law as well as the background field alone in the source free field equations, without introducing any other arbitrary ingredients or modification of the General Theory of Relativity.

Submitted on: May 04, 2013 / Accepted on: May 18, 2013

References

1. Einstein A. The meaning of Relativity. Appendix II. Translated in French by M. Solovine and M. A. Tonnelat, Gauthier-Villars, Paris, 1954.
2. Tonnelat M. A. Les Théories Unitaires de l'Electromagnétisme et de la Gravitation. Gauthier-Villars, Paris, 1959, p.18.
3. Eddington A. S. The mathematical Theory of Relativity. Cambridge University Press, Cambridge, 1924.
4. Lichnérowicz A. Les Théories Relativistes de la Gravitation et de l'Electromagnétisme. Masson et Cie, Paris, 1955.
5. Cartan E. La Géométrie des Espaces de Riemann. 1925. Re-print, Gauthier-Villars, Paris, 1946.
6. Landau L., Lifshitz E. The Classical Theory of Fields. Addison-Wesley, Reading, Massachusetts, 1962.
7. Straumann N. General Relativity and Relativistic Astrophysics. Springer-Verlag, Berlin, 1984, p.159.
8. Hawking S.W., Ellis G.F.R. The Large Scale Structure of Space-Time. Cambridge University Press, Cambridge, 1987, p.62.
9. Dirac P.A.M. General Theory of Relativity. Princeton University Press, 2nd edition, 1975, p.61.
10. Tolman R.C. Static Solutions of Einstein's Field Equations for Sphere of Fluid. *Physical Review*, 1939, v. 55, no. 4, 364–373.
11. Kramer D., Stephani H., Hertl E., Mac Callum M. Exact Solutions of Einstein's Field Equations. Cambridge University Press, Cambridge, 1979.

The Role of Evection in Optical Measurements of Light Beam Deflection from the Sun's Disk (the Einstein Effect)

Sergey N. Shapovalov

SCC RF Arctic and Antarctic Research Institute. 38 Bering St., St. Petersburg 199397, Russia
E-mail: shapovalov@aari.ru, tel. +7 (812) 3373157

The relationship between the optical results of light beam deflection from the disk of the Sun ($\delta\varphi$) obtained during observations of the total solar eclipses, from 1919 till 1973, and the evection, the major perturbation from the Sun, based on the theory of the Moon's motion, is analysed. The dependence of $\delta\varphi$ upon the temporal changes of the evection was found. The expected $\delta\varphi$ optical results for the total solar eclipses, for the period from 22.09.2003 till 29.12.2103, were calculated. Based on the comparison of calculated evection values with fluctuations of intensity of solar radiation within 603–607 nm range obtained through the spectral observations on solar radiation in Antarctica, the modulatory role of the evection in deflecting the light beam at the near-Earth space was concluded.

Optical measurements of the star beam deflection from the Sun disk were performed by a number of researchers during the total solar eclipses, from 29.05.1919 till 30.06.1973, with the purpose of checking the $\delta\varphi$ angle value (1.75″) obtained by Einstein, following his development of the General Theory of Relativity (GTR) [1]. In case the radio measurements only are considered in the practical estimates of the Einstein effect, $\delta\varphi$ values match with the theory within 1% range [2]. For example, an average value of 1.73″(±0.07″) was obtained in radar measurements of Mercury, Venus and Mars, whereas measurements of quasars and pulsars using radio interferometry produced an estimate of 1.76″(±0.08). Deflection of the beam from the Sun disk is described by the equation:

$$\delta\varphi = -\frac{4GM_{\odot}}{R_{\odot}c^2}\,, \qquad (1)$$

where the "minus" sign corresponds to the deflection of the beam to the center of the Sun; $G = 6.67{\times}10^{-11}$ H×m^2/kg^2 is the gravitational constant; $M_{\odot} = 1.99{\times}10^{30}$ kg is the mass of the Sun; $c = 3{\times}10^8$ m/s is the speed of light; $R_{\odot} = 6.96{\times}10^8$ m is the radius of the Sun.

Based on the optical observations of the eight total solar eclipses, the author's average result together with a confidence interval of measurements makes $\delta\varphi = 1.83 \pm 0.40$, and the recalculated measurement result is $\delta\varphi = 2.0 \pm 0.13$, which, in view of the low accuracy and the considerable spread of measurements, is consistent with the GTR. According to the published data [3–10], the results of $\delta\varphi$ optical measurements for the total solar eclipses observed from 1919 till 1973 were as follows:

29.05.1919: (1.98, 0.93, 1.61),

21.09.1922: (1.42, 1.75, 2.16, 1.72, 1.83, 1.77),

09.05.1929: (2.24),

19.06.1936: (2.73, 2.13, 1.28),

20.05.1947: (2.01),

25.02.1952: (1.70, 1.82),

02.10.1959: (2.17),

30.06.1973: (1.66).

Observations referring to the date 19.06.1936 should be considered as ineffectual, since the absolute value error exceeds 200%. To date, the list of known errors includes:

— Deviation of the Sun's shape from the sphericity, $9.2'' {\times} 10^{-2}$;

— The Earth's motion along the ecliptic ($2.88'' {\times} 10^{-2}$);

— Beam refraction in the atmosphere of the Sun (0.004″);

— Refraction and dispersion in the Earth atmosphere (0.01″–0.1″);

— Offset of the observer from the Sun-Moon-Earth line;

— The influence of the gravitational field of the Moon and the Earth during the total eclipse event, by an addition to the relativistic beam deflection ($5.8'' {\times} 10^{-4}$);

— Wavelength dependence of the light beam ($2.5'' {\times} 10^{-4}$);

— Dependence on solar activity;

— Astroclimatic characteristics of a particular observation station;

— Additive error caused by inaccurate scale matching between the day and night astroimages (0.25″).

It should be noted that through the history of $\delta\varphi$ measurements the list of errors has expanded considerably; however, the accuracy of estimates is not yet improved. Summing the values of all the errors, the magnitude of the total correction is apparently insignificant. Therefore, dispersion of $\delta\varphi$ results is probably due to the influence of some unknown factors.

The major solar-induced disturbances are described by terms in the formula of the geocentric ecliptic longitude of the Moon [11, 12]. Full description of this formula includes 1,500 terms [13], where evection, variation and annual inequality are the most important. When limited to the largest

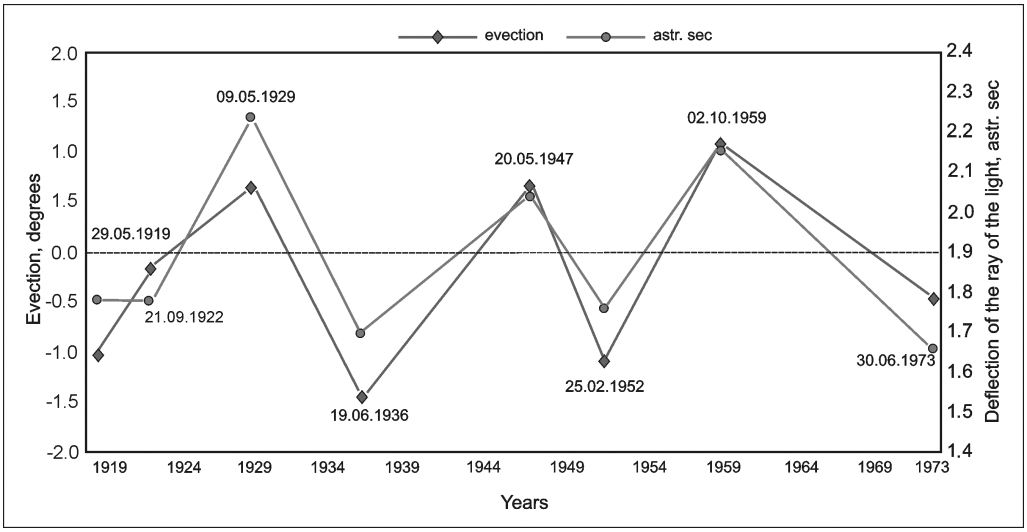

Fig. 1: Comparison of the evection angle values with the results of optical measurements taken as average values and excluding errors for the dates of the total solar eclipses, 1919–1973.

in amplitude terms, the formula is as follows:

$$\lambda = L + 6.289^\circ \sin l - 1.274^\circ \sin(l - 2D) + \\ + 0.658^\circ \sin 2t + 0.214^\circ \sin 2l - \\ - 0.186^\circ \sin l' - 0.114^\circ \sin 2F\,, \qquad (2)$$

where L is the mean longitude (void of the periodic disturbances) of the Moon in the orbit, l, D, l', F are the main arguments in the lunar theory.

In the first five inequalities of the formula (2), the terms bearing coefficients 6.289 and 0.214 are determined by ellipticity of unperturbed (Keplerian) orbit, whereas the terms with coefficients 1.274 (evection, 31.8 days), 0.658 (variation, 14.8 days) and 0.186 (annual inequality, 186.2 days) are caused by gravitational perturbations from the Sun. The periods of these inequalities, according to the theory of motion of the Moon, exist in the short-period nutation of the Earth's axis, as well [14]. In this paper we consider the contribution of the evection, the main and the largest in amplitude perturbation from the Sun, as the most significant deviation of the true motion of the Moon from its motion defined by Kepler's laws. Evection was discovered by Ptolemy (2AD) when observing the Moon in the 1st and 3rd quarters (in quadrature points). The physical explanation of the evection was developed by Newton. Evection can be represented as a difference in the equation of the center [13] generated by the term $1.274^\circ \sin(l - 2D)$:

$$e_\odot = 5.02 \sin l + 0.214 \sin 2l\,, \qquad (3)$$

$$e_\odot = 7.56 \sin l + 0.214 \sin 2l\,. \qquad (4)$$

This effect is determined by the gravitational influence of the Sun to the Moon. In syzygial points of the lunar orbit (new moon and full moon), this term is subtracted from the senior term of the equation (3), and it is added in quadrature. During the new moon and full moon, $2D = 0^\circ$, or 360° (3), which is the same in the context of trigonometric functions. In the first and last quarters, $D = 90^\circ$, or 270° (4). So, the known manifestations of the evection in the near-Earth space motivated the studies of its contribution to the results of $\delta\varphi$ assessments obtained during observations of the total solar eclipses, from 1919 till 1973.

The *evection* values were calculated upon the Julian dates of the total solar eclipses. Fig. 1 shows a comparison of the evection angle values with the results of optical measurements taken as average values and excluding errors for the dates of the total solar eclipses. Anomalous results 0.93″ (1919) and 2.73″ (1936) were omitted from the calculations of average values, as they fell outside the range of average result and the confidence interval of all measurements.

Fig. 2 shows the distribution of dependency of optical results from the evection. Continuous curve, which includes 0.93″ (1919) and 2.73″ (1936) values, represents averaging of results depending on the evection and is described as:

$$\delta\varphi(M) = 1.7227 + 0.2058\,x + 0.3163\,x^2. \qquad (5)$$

The dotted curve, which excludes 0.93″ (1919) and 2.73″ (1936) values, represents averaging of results depending on the evection and is described as follows:

$$\delta\varphi(E) = 1.723 + 0.316\,x^2. \qquad (6)$$

As demonstrated in the Figure, $\delta\varphi(M)$ has a lower left-hand shift against $\delta\varphi(E)$ characterized by the term $0.2058\,x$ (5), due to the low values obtained during the observations of 1919 ($\delta\varphi = 0.93''$) and 1936 ($\delta\varphi = 1.28''$). According to $\delta\varphi(E)$ distribution in Fig. 2, deflection of beams in

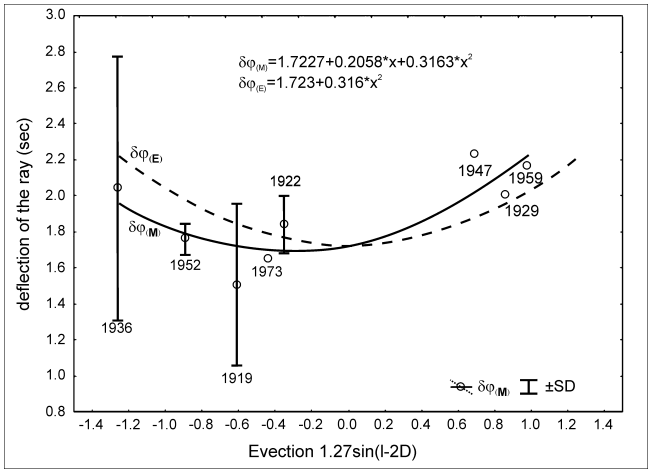

Fig. 2: Distribution of $\delta\varphi$ values according to the *evection* values, 1919–1973: $\delta\varphi(M)$ averaged optical results dependent on the evection, 0.93″ (1919) and 2.73″ (1936) values included; $\delta\varphi(E)$ averaged optical results dependent on the evection, 0.93″ (1919) and 2.73″ (1936) values excluded.

the evection extremes ($\pm 1.274^\circ$) should correspond to $\delta\varphi \approx 2.25 \pm 10\%$, and to $\delta\varphi \approx 1.72$ in case of 0, i.e., conform to the Einstein result. Using the expression (6), the expected $\delta\varphi$ values calculated for optical observations at the dates of the total solar eclipse, from 23.11.2003 till 29.12.2103, are presented in Table 1.

Along with the deviations in the motion of the Moon from the Keplerian orbit and the short-period nutation of the Earth axis, the evection mechanism is detected in spectral zenith observations of the atmosphere at Novolazarevskaya station (Antarctica). These observations are aimed to investigate the fluctuations of energy and intensity of scattered solar UV radiation under the 11-year SA cycle. Measurements of fluctuations are recorded in the following ranges: 303–305 nm, 331–332.5 nm, 329.5–334 nm, 336–345 nm, 297–307 nm, 321–331 nm, 297–330 nm, and 603–607 nm, during the polar summer (September – February). Detailed description of the methodology of observations is cited in [15].

To test the influence of the evection factor on variations of the light flux, fluctuations measurements in the range of 603–607 nm (as the most proximate band to the central part of the solar spectrum) were selected from the available set of registered channels. Based on the observations during the polar summer 2007–2008 and 2008–2009, data analysis of the intensity channel was performed, in average daily standard deviation (SD) units, to build the time series and provide temporal comparison with the calculated values of the evection. Figs. 3 and 4 show the distribution pattern of SD values (603–607 nm), to be compared with the evection changes.

The figures show a reasonably good phase and periodic matching between the SD (603–607 nm) dynamics and the evection changes during the polar summer of 2007–2008. However, Fig. 4 shows the broken phase matching at certain

Eclipses	$\delta\varphi(E)$	Eclipses	$\delta\varphi(E)$	Eclipses	$\delta\varphi(E)$
23.11.2003	1.75	26.12.2038	2.02	03.08.2073	2.11
08.04.2005	2.24	21.06.2039	1.82	27.01.2074	2.23
03.10.2005	2.09	15.12.2039	1.72	24.07.2074	2.13
29.03.2006	1.89	30.04.2041	2.23	16.01.2075	2.03
22.09.2006	1.74	25.10.2041	2.01	13.07.2075	1.76
07.02.2008	2.22	20.04.2042	1.81	06.01.2076	1.73
01.08.2008	2.16	14.10.2042	1.72	22.05.2077	2.19
26.01.2009	1.88	28.02.2044	2.22	15.11.2077	2.02
22.07.2009	1.74	23.08.2044	2.09	11.05.2078	1.83
15.01.2010	1.75	16.02.2045	1.89	04.11.2078	1.72
11.07.2010	1.9	12.08.2045	1.74	01.05.2079	1.86
20.05.2012	1.78	05.02.2046	1.74	24.10.2079	1.96
13.11.2012	1.72	02.08.2046	1.98	10.03.2081	1.82
10.05.2013	1.91	11.06.2048	1.74	03.09.2081	1.72
03.11.2013	2.11	05.12.2048	1.75	27.02.2082	1.79
09.03.2016	1.92	31.05.2049	1.89	24.08.2082	2.07
01.09.2016	2.02	25.11.2049	2.17	03.07.2084	1.72
26.02.2017	2.23	20.05.2050	2.22	27.12.2084	1.79
21.08.2017	2.23	30.03.2052	1.91	22.06.2085	1.98
02.07.2019	2.03	22.09.2052	2.01	16.12.2085	2.16
26.12.2019	2.2	20.03.2053	2.23	11.06.2086	2.24
21.06.2020	2.18	12.09.2053	2.2	21.04.2088	1.99
14.12.2020	2.11	24.07.2055	2.11	14.10.2088	2.09
10.06.2021	1.82	16.01.2056	2.19	10.04.2089	2.24
04.12.2021	1.72	12.07.2056	2.19	04.10.2089	2.14
20.04.2023	2.22	05.01.2057	2.02	23.09.2090	1.77
14.10.2023	2.1	01.07.2057	1.83	15.08.2091	2.18
08.04.2024	1.89	26.12.2057	1.73	07.02.2092	2.23
02.10.2024	1.75	11.05.2059	2.23	03.08.2092	2.13
17.02.2026	2.22	05.11.2059	2.01	27.01.2093	1.94
12.08.2026	2.17	30.04.2060	1.82	23.07.2093	1.77
06.02.2027	1.88	24.10.2060	1.72	16.01.2094	1.73
02.08.2027	1.74	20.04.2061	1.86	02.06.2095	2.19
26.01.2028	1.74	13.10.2061	1.96	27.11.2095	1.93
22.07.2028	1.99	28.02.2063	1.81	22.05.2096	1.76
01.06.2030	1.74	24.08.2063	1.72	15.11.2096	1.73
25.11.2030	1.75	17.02.2064	1.79	11.05.2097	1.85
21.05.2031	1.9	12.08.2064	1.98	04.11.2097	2.05
14.11.2031	2.1	22.06.2066	1.74	21.03.2099	1.82
09.05.2032	2.23	17.12.2066	1.8	14.09.2099	1.72
30.03.2033	1.78	11.06.2067	1.89	10.03.2100	1.78
20.03.2034	1.91	06.12.2067	2.17	04.09.2100	2.06
12.09.2034	2.02	31.05.2068	2.24	28.02.2101	2.21
09.03.2035	2.23	11.04.2070	2	15.07.2102	1.72
02.09.2035	2.19	04.10.2070	2.1	08.01.2103	1.79
13.07.2037	2.12	31.03.2071	2.23	04.07.2103	1.97
05.01.2038	2.19	23.09.2071	2.2	29.12.2103	2.22
02.07.2038	2.19	12.09.2072	1.77		

Table 1: Expected $\delta\varphi$ results for the total solar eclipses, from 23.11.2003 till 12.29.2103.

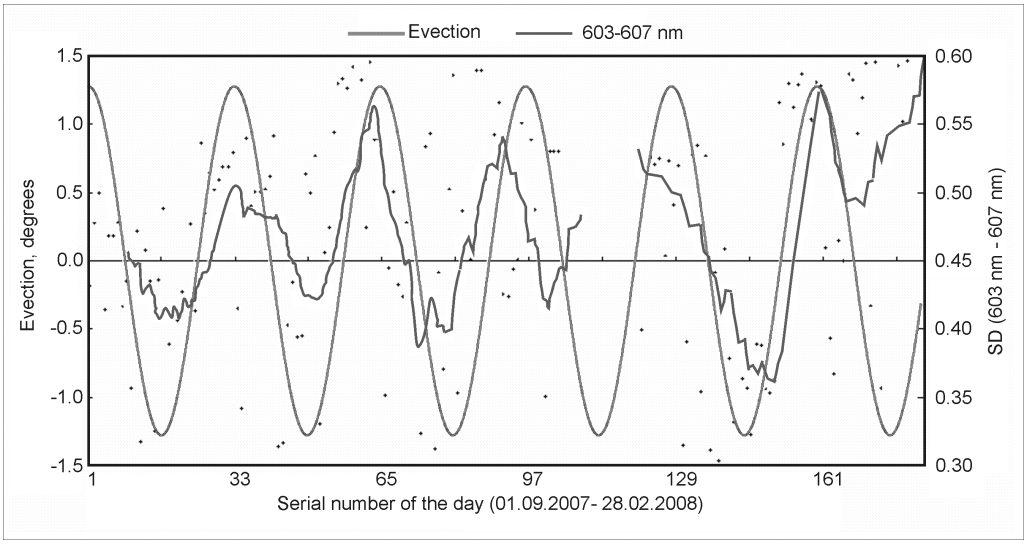

Fig. 3: Comparison of temporal changes in the evection and the average daily standard deviation (SD) of radiation intensity in the 603–607 nm (9 pt. mov. aver.) range, for the period from 01.09.2007 till 28.02.2008.

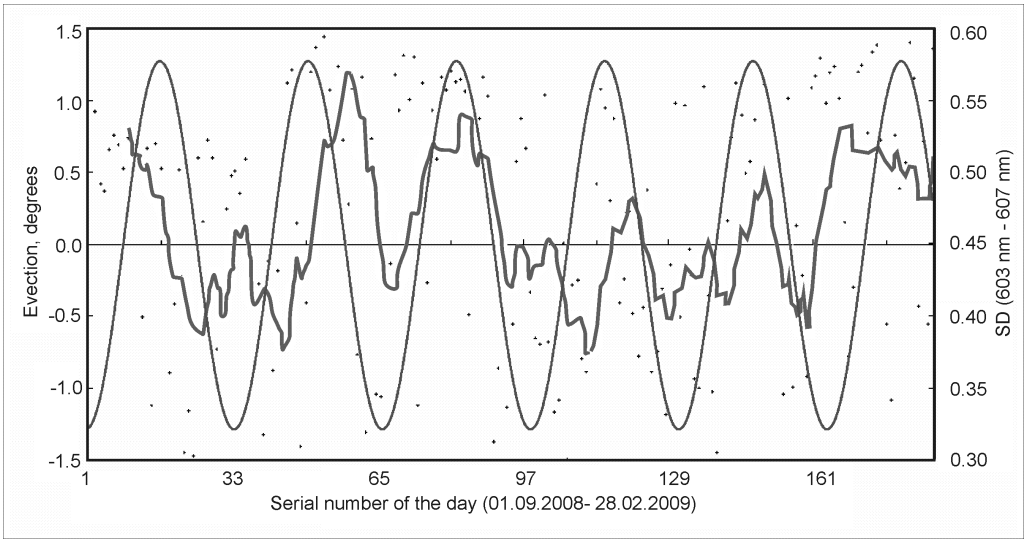

Fig. 4: Comparison of temporal changes in the evection and the average daily standard deviation (SD) of radiation intensity in the 603–607 nm (9 pt. mov. aver.) range, for the period from 01.09.2008 till 28.02.2009.

extended sections. In our view, such failures may be related to the SA stages. Among the above errors, $\delta\varphi$ dependence from SA and astroclimatic characteristics of observation stations remain understudied. Astroclimatic characteristics are determined by the weather conditions and optical properties of the atmosphere and both are connected with the SA manifestations. Although the mechanism of SA effects on the surface layer of the atmosphere remains unclear to date, this connection is revealed by the long-term observations of weather services.

In a brief discussion of relationship between $\delta\varphi$ and the evection, previously disregarded in research practice, a 3-body Einstein model should be mentioned, which considers the Earth and the Moon as point-like objects. This model is undeniable in the evaluation of mass gravitation of the Earth–Moon and the Sun. The major solar disturbances cause deviation from the Keplerian orbit of the Moon motion and, at the same time, deviations in the Earth axis in the short-period nutation (31.8 and 14.8 days), provide periodic gravitational influence on the Earth–Moon system. Obviously, this influence

is manifested in the Einstein effect through the modulation property of optical beams.

Conclusions

— The values of $\delta\varphi$ optical results reveal statistical correlation with the temporal change of the evection;

— In the evection extreme points ($\pm 1.274^\circ$), deflection of optical beams from the solar disk is expected to approach $\delta\varphi \approx 2.25 \pm 10\%$;

— When the evection values $\approx 0^\circ$, it is expected to approach $\delta\varphi = 1.72 \pm 10\%$;

— In conformity with $\delta\varphi(E)$, introduction of correction for the evection into the formula (1) is justified.

Submitted on: April 26, 2013 / Accepted on: May 06, 2013

References

1. Einstein A. *Ann. Phys.*, 1916, Bd.49, 769.
2. Ginzburg V.L. Theoretical physics and astrophysics. Nauka, Moscow, 1980.
3. Crommelin A. Results of the total solar eclipse of May 29 end the relativity theories. *Nature*, 1920, v.104, 280–281.
4. Hopmann J. Die Deutung der Ergebnisse der amerikanischen Einstein expedition. *Phys. Z.*, 1923, Bd.24, No.21/22, 476–485.
5. Eddington A.S. The deflection of light during a solar eclipse. *Nature*, 1919, v.104, no.271, 454.
6. Mikhailov A.A. Einstein's effect observations. *Astronomicheskiy Zhurnal*, 1956, no.33, 912–927.
7. Freundlich E.F., von Kluber H.V., Brunn A. Ergebniss der Potsdamer Expedition zur Beib chtung der Sonnenfinsternis von 1929, Mai 9 die Ablenkung des Lichtes in Schwerefeld der Sonne. *Phys. Ber.*, 1931, No.1, 2838–2839.
8. Schmeidler F. Neuer Versuch einer Messung der relativitischen Lichtablenkung. *Astron. Nachrichten*, 1963, Bd.287, No.1-2, 7–16.
9. Jones B.F. Texas Mauritanian Eclipse Team. Gravitational deflection of light: Solar eclipse of 30 June 1973. I. Description of procedures and final results. *Astron. J.*, 1974, v.81, 452–454.
10. Jones B.F. Gravitational deflection of light: solar eclipse of 30 June 1973. II. Plate reduction. *Astron. J.*, 1974, v.81, 455–463.
11. Astronomicheskiy Ezhegodnik. Part 1. Inst. Appl. Astron., Russian Acad. Sci., St. Petersburg, 1998.
12. Meeus J. Astronomical formulae for calculators. Mir, Moscow, 1988.
13. Brown F.H. Tables of the motion of the Moon. New Hawen, 1919.
14. Kulikov K.A. Rotation of the Earth. Nedra, Moscow, 1985.
15. http://www.aari.nw.ru/clgmi/geophys/data_nvl_ru.html

Mass and Charge Selfvariation: A Common Underlying Cause for Quantum Phenomena and Cosmological Data

Emmanuil Manousos

Astrophysics Laboratory, Faculty of Physics, National and Kapodistrian University of Athens, Panepistimiopolis, GR 15783 Zographos, Athens, Greece. E-mail: emanoussos@phys.uoa.gr

The physical theories of the last century do not possess the completeness necessary in order to justify the quantum phenomena and the cosmological data. In this article, we present the law of selfvariations and suggest it as the common cause of quantum and cosmological phenomena. There is an intermediate state between matter and the photon, which is the cause of quantum phenomena. The cosmological data are condensed in a single equation with one unknown. The consequences of the law of selfvariations extend from the microcosm to the observations we make billions of light-years away.

Contents:

1 Introduction

The study we present in the current edition is based on two assumptions that are taken as axioms. The first assumption is that the rest masses m_0 and electric charges q of material particles increase with the passage of time (selfvariations). The second assumption is that the consequences of the selfvariations propagate through four-dimensional spacetime with a zero arc length: $dS^2 = 0$. The set of consequences arising from these two assumptions constitutes the "theory of selfvariations".

An immediate consequence of the statements-axioms we have introduced, is the concept of the generalized photon: a particle carrying energy E, linear momentum $\mathbf{P}$, and moving with velocity $\boldsymbol{\upsilon}$, of magnitude $\|\boldsymbol{\upsilon}\| = c$, in every inertial frame of reference. The generalized photon correlates the material particle with its surrounding spacetime. In its simplest version, the generalized photon is emitted by the material particle into its surrounding spacetime. When the material particle is electrically charged, the generalized photon, apart from energy and momentum, also carries electric charge.

In figure 1, the arbitrary motion of a material point particle moving with velocity $\mathbf{u}$ in an inertial frame of reference $O(x, y, z, t)$ is represented.

A generalized photon is emitted by the material particle at time $w = t - \frac{r}{c}$, from point $E(x_p(w), y_p(w), z_p(w), w)$, and arrives at time t at point $A(x, y, z, t)$. The velocity of the generalized photon in Figure 1, is

$$\boldsymbol{\upsilon} = \frac{c}{r}\mathbf{r}$$

where $r = \|\mathbf{r}\|$. We express the vector $\frac{\boldsymbol{\upsilon}}{c}$ in the trigonometric form

$$\frac{\boldsymbol{\upsilon}}{c} = \begin{bmatrix} \frac{\upsilon_x}{c} \\ \frac{\upsilon_y}{c} \\ \frac{\upsilon_z}{c} \end{bmatrix} = \begin{bmatrix} \cos\delta \\ \sin\delta\cos\omega \\ \sin\delta\sin\omega \end{bmatrix}.$$

Furthermore, we define the following two vectors

$$\boldsymbol{\beta} = \begin{bmatrix} -\sin\delta \\ \cos\delta\cos\omega \\ \cos\delta\sin\omega \end{bmatrix}$$

and

$$\boldsymbol{\gamma} = \begin{bmatrix} 0 \\ -\sin\omega \\ \cos\omega \end{bmatrix}.$$

The vectors $\frac{\boldsymbol{\upsilon}}{c}, \boldsymbol{\beta}, \boldsymbol{\gamma}$ constitute a right-handed, orthonormal vector basis that accompanies the generalized photon in

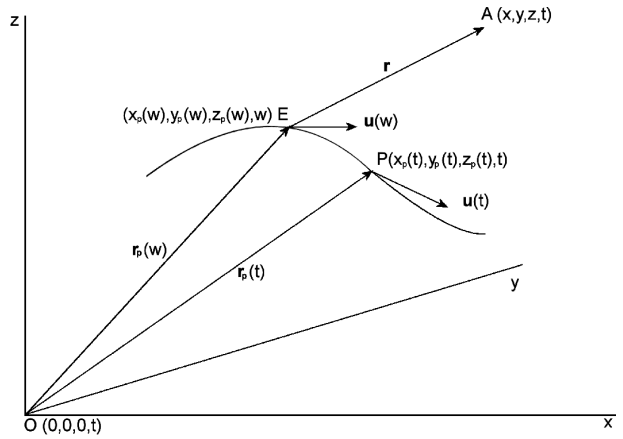

Fig. 1: A material point particle moving arbitrarily. As the material particle moves from point $E(x_p(w), y_p(w), z_p(w), w)$ to point $P(x_p(t), y_p(t), z_p(t), t)$, the generalized photon moves from point $E(x_p(w), y_p(w), z_p(w), w)$ to point $A(x, y, z, t)$.

its motion. The consequences of the selfvariations are expressed as functions of the parameters $w = t - \frac{r}{c}, r, \delta, \omega$. The basic study of the selfvariations leads to two fundamental theorems: the "Fundamental Mathematical Theorem", and the "Trajectory Representation Theorem". The first theorem allows us to correlate any change in energy manifested on the material particle at point $E(x_p(w), y_p(w), z_p(w), w)$ with a corresponding change in energy at point $A(x, y, z, t)$ of Figure 1. The second theorem represents the tangent vector, the curvature and the torsion of the trajectory of the material particle onto the geometric characteristics of the generalized photon in the surrounding spacetime. The two theorems allow us to express quantitatively the consequences of the selfvariations on the surrounding spacetime of the material particle. As a consequence of the selfvariations, in the surrounding spacetime of the material particle there is energy of density D

$$D = -c\frac{\partial m_0}{\partial w}\frac{1}{4\pi\gamma^3 r^2\left(1-\dfrac{\boldsymbol{\upsilon}\cdot\mathbf{u}}{c^2}\right)^4}$$

and momentum of density $\mathbf{J}$

$$\mathbf{J} = D\frac{\boldsymbol{\upsilon}}{c^2}$$

where

$$\gamma = \frac{1}{\sqrt{1-\dfrac{u^2}{c^2}}},$$

and $\mathbf{u} = \mathbf{u}(w)$.

If the material particle is electrically charged, then in the surrounding spacetime there is also electric charge of density ρ

$$\rho = -\frac{\partial q}{c\partial w}\frac{1}{4\pi\gamma^2 r^2\left(1-\dfrac{\boldsymbol{\upsilon}\cdot\mathbf{u}}{c^2}\right)^3}$$

and electric current of density $\mathbf{j}$

$$\mathbf{j} = \rho\boldsymbol{\upsilon}.$$

The Lienard-Wiechert potentials

$$V = \frac{q}{4\pi\varepsilon_0 r\left(1-\dfrac{\boldsymbol{\upsilon}\cdot\mathbf{u}}{c^2}\right)}$$

and

$$\mathbf{A} = \frac{q}{4\pi\varepsilon_0 c^2 r\left(1-\dfrac{\boldsymbol{\upsilon}\cdot\mathbf{u}}{c^2}\right)}\mathbf{u}$$

are not compatible with the theory of selfvariations. Therefore, they are replaced by the potentials of the selfvariations

$$V = \frac{q\left(1-\dfrac{u^2}{c^2}\right)}{4\pi\varepsilon_0 r\left(1-\dfrac{\boldsymbol{\upsilon}\cdot\mathbf{u}}{c^2}\right)^2} + \frac{q\,(\boldsymbol{\upsilon}\cdot\boldsymbol{\alpha})}{4\pi\varepsilon_0 c^3\left(1-\dfrac{\boldsymbol{\upsilon}\cdot\mathbf{u}}{c^2}\right)^2}$$

$$\boldsymbol{A} = V\frac{\boldsymbol{\upsilon}}{c^2}$$

where $\boldsymbol{\alpha} = \boldsymbol{\alpha}(w)$ is the acceleration of the material particle.

The potentials of the selfvariations are separated into two individual pairs

$$V_u = \frac{q\left(1-\dfrac{u^2}{c^2}\right)}{4\pi\varepsilon_0 r\left(1-\dfrac{\boldsymbol{\upsilon}\cdot\mathbf{u}}{c^2}\right)^2}$$

$$\boldsymbol{A}_u = V_u\frac{\boldsymbol{\upsilon}}{c^2}$$

and

$$V_\alpha = \frac{q\,(\boldsymbol{\upsilon}\cdot\boldsymbol{\alpha})}{4\pi\varepsilon_0 c^3\left(1-\dfrac{\boldsymbol{\upsilon}\cdot\mathbf{u}}{c^2}\right)^2}$$

$$\boldsymbol{A}_\alpha = V_\alpha\frac{\boldsymbol{\upsilon}}{c^2}.$$

The $(V_u, \boldsymbol{A}_u)$ pair gives the electromagnetic field $(\boldsymbol{\varepsilon}_u, \boldsymbol{B}_u)$ that accompanies the electrically charged material particle

$$\boldsymbol{\varepsilon}_u = \frac{q\left(1-\dfrac{u^2}{c^2}\right)}{4\pi\varepsilon_0 r^2\left(1-\dfrac{\boldsymbol{\upsilon}\cdot\mathbf{u}}{c^2}\right)^3}\left(\frac{\boldsymbol{\upsilon}}{c}-\frac{\mathbf{u}}{c}\right)$$

$$\boldsymbol{B}_u = \frac{q\left(1-\dfrac{u^2}{c^2}\right)}{4\pi\varepsilon_0 r^2\left(1-\dfrac{\boldsymbol{\upsilon}\cdot\mathbf{u}}{c^2}\right)^3}\frac{\mathbf{u}}{c}\times\frac{\boldsymbol{\upsilon}}{c}.$$

The $(V_\alpha, \boldsymbol{A}_\alpha)$ pair gives the electromagnetic radiation

$$\boldsymbol{\varepsilon}_\alpha = \frac{q}{4\pi\varepsilon_0 c^2 r\left(1-\dfrac{\boldsymbol{\upsilon}\cdot\mathbf{u}}{c^2}\right)^2}\left[\frac{\dfrac{\boldsymbol{\upsilon}}{c}\boldsymbol{\alpha}}{1-\dfrac{\boldsymbol{\upsilon}\cdot\mathbf{u}}{c^2}}\left(\frac{\boldsymbol{\upsilon}}{c}-\frac{\mathbf{u}}{c}\right)-\boldsymbol{\alpha}\right]$$

$$\boldsymbol{B}_\alpha = \frac{q}{4\pi\varepsilon_0 r\left(1-\dfrac{\boldsymbol{\upsilon}\cdot\mathbf{u}}{c^2}\right)}\left[\frac{\frac{\upsilon}{c}\boldsymbol{\alpha}}{1-\dfrac{\boldsymbol{\upsilon}\cdot\mathbf{u}}{c^2}}\left(\frac{\mathbf{u}}{c}\times\frac{\boldsymbol{\upsilon}}{c}\right)-\frac{\boldsymbol{\upsilon}}{c}\times\boldsymbol{\alpha}\right].$$

The pair $(V_\alpha, \boldsymbol{A}_\alpha)$ of the electromagnetic radiation potentials does not depend on the distance r. For each couple $(\boldsymbol{\varepsilon}, \boldsymbol{B})$ the following relation holds

$$\boldsymbol{B} = \frac{\boldsymbol{\upsilon}}{c^2}\times\boldsymbol{\varepsilon}.$$

The energy-momentum tensor for the generalized photon that results from the selfvariation of the rest mass m_0 of the material particle is given by the matrix Φ^{ij}

$$\Phi^{ij} = \frac{D}{c^2}\begin{bmatrix} c^2 & c\upsilon_x & c\upsilon_y & c\upsilon_z \\ \upsilon_x c & \upsilon_x^2 & \upsilon_x\upsilon_y & \upsilon_x\upsilon_z \\ \upsilon_y c & \upsilon_y\upsilon_x & \upsilon_y^2 & \upsilon_y\upsilon_z \\ \upsilon_z c & \upsilon_z\upsilon_x & \upsilon_z\upsilon_y & \upsilon_z^2 \end{bmatrix}$$

where

$$\begin{bmatrix} c \\ \upsilon_x \\ \upsilon_y \\ \upsilon_z \end{bmatrix} = \begin{bmatrix} \upsilon^0 \\ \upsilon^1 \\ \upsilon^2 \\ \upsilon^3 \end{bmatrix}, \qquad i, j = 0, 1, 2, 3.$$

The energy-momentum tensor for the generalized photon that results from the selfvariation of the electric charge q of the material particle is given by the matrix Φ^{ij}

$$\Phi^{ij} = \begin{bmatrix} W & cS_x & cS_y & cS_z \\ cS_x & \sigma_{11} & \sigma_{12} & \sigma_{13} \\ cS_y & \sigma_{21} & \sigma_{22} & \sigma_{23} \\ cS_z & \sigma_{31} & \sigma_{32} & \sigma_{33} \end{bmatrix} - \frac{\rho V}{c^2}\begin{bmatrix} c^2 & c\upsilon_x & c\upsilon_y & c\upsilon_z \\ \upsilon_x c & \upsilon_x^2 & \upsilon_x\upsilon_y & \upsilon_x\upsilon_z \\ \upsilon_y c & \upsilon_y\upsilon_x & \upsilon_y^2 & \upsilon_y\upsilon_z \\ \upsilon_z c & \upsilon_z\upsilon_x & \upsilon_z\upsilon_y & \upsilon_z^2 \end{bmatrix}$$

where $(S_x, S_y, S_z) = \mathbf{S} = \varepsilon_0\boldsymbol{\varepsilon}\times\boldsymbol{B}$ is the Poynting vector,

$$W = \frac{1}{2}\varepsilon_0\left(\boldsymbol{\varepsilon}^2 + c^2\boldsymbol{B}^2\right)$$

and

$$\sigma_{\alpha\beta} = \varepsilon_0\left(-\varepsilon_\alpha\varepsilon_\beta - c^2 B_\alpha B_\beta + W\delta_{\alpha\beta}\right)$$

$$\delta_{\alpha\beta} = \begin{cases} 1, & \text{if } \alpha = \beta \\ 0, & \text{if } \alpha \neq \beta \end{cases}$$

where $\alpha, \beta = 1, 2, 3$ and

$$(\varepsilon_1, \varepsilon_2, \varepsilon_3) = (\varepsilon_x, \varepsilon_y, \varepsilon_z) = \boldsymbol{\varepsilon}$$

$$(B_1, B_2, B_3) = (B_x, B_y, B_z) = \boldsymbol{B}.$$

The energy-momentum tensors Φ^{ij} give us important information about the energy content of the surrounding spacetime of the material particle. Furthermore, they are related with the gravitational and the electromagnetic interaction. As we progress in our study however, it becomes evident that there is information about the energy content and the properties of spacetime, that is not contained within the Φ^{ij} tensors.

The study we presented up to this point has been conducted without a quantitative determination of the selfvariations. We made the assumption of the selfvariations in order to undertake the relevant calculations, but we have not determined quantitatively the rate at which they evolve, i.e. the $\frac{\partial m_0}{\partial w}$ and $\frac{\partial q}{\partial w}$. In order to study the consequences of the selfvariations, we have to quantitatively determine these rates.

The quantitative determination of the selfvariations is made on the basis of the total energy E_s and the total momentum $\mathbf{P_s}$ emitted simultaneously in all directions, by the material particle. The rest mass m_0 and the electric charge q of the material particle vary according to the operators

$$\frac{\partial}{\partial t} \to -\frac{i}{\hbar}E_s$$

$$\nabla \to \frac{i}{\hbar}\mathbf{P_s}$$

where h is Planck's constant, and $\hbar = \frac{h}{2\pi}$. The law of selfvariations expresses a continuous interaction between the material particle and the generalized photons.

The partial contribution of an individual generalized photon to the law of selfvariations is determined by the percentage-function Φ. Due to this, function Φ has a fundamental role in the energy content of the generalized photon.

The energy E and momentum $\mathbf{P}$ of the generalized photon that is related to the selfvariation of the rest mass m_0 of the material particle, are given by the equations

$$E = \Phi\frac{i\hbar}{1-\dfrac{\boldsymbol{\upsilon}\cdot\mathbf{u}}{c^2}}\frac{\partial m_0}{m_0\partial w} + i\hbar\frac{\partial\Phi}{\partial t}$$

$$\mathbf{P} = \Phi\frac{i\hbar}{1-\dfrac{\boldsymbol{\upsilon}\cdot\mathbf{u}}{c^2}}\frac{\partial m_0}{m_0 c\partial w}\frac{\boldsymbol{\upsilon}}{c} - i\hbar\nabla\Phi.$$

The equations that give the energy and momentum of the generalized photon that is related to the selfvariation of the electric charge of the material particle, are of similar form.

The energy E and the momentum $\mathbf{P}$ of the generalized photon do not obey the simple relation

$$\mathbf{P} = E\frac{\boldsymbol{\upsilon}}{c^2}.$$

That relation is a special case of the general relation

$$\mathbf{P} = E\frac{\boldsymbol{\upsilon}}{c^2} - \frac{i\hbar}{r}\frac{\partial\Phi}{\partial\delta}\boldsymbol{\beta} - \frac{i\hbar}{r\sin\delta}\frac{\partial\Phi}{\partial\omega}\boldsymbol{\gamma}.$$

The generalized photon determines the relation of the material particle with the surrounding spacetime. Furthermore, it is related with the energy content of spacetime and, hence, with the very properties of spacetime. Because of this, a large part of the study we present in the present edition concerns the generalized photon and its properties. The resulting equations contain an exceptionally large body of data and information.

Thus, we shall confine ourselves to a brief report for the structure and the properties of the generalized photon.

The generalized photon carries four energy-momentum pairs, each of which transforms autonomously, independently of the rest, according to Lorentz-Einstein. Two of these pairs do not possess rest energy, do not depend on the distance r from the material particle, are defined both on the material particle and on the surrounding spacetime, while they do not possess intrinsic angular momentum (spin). The other two energy-momentum pairs have, respectively, rest energy

$$\pm\frac{c\hbar}{r}\frac{\partial\Phi}{\partial\delta}$$

$$\pm\frac{c\hbar}{r\sin\delta}\frac{\partial\Phi}{\partial\omega}.$$

Their energy and momentum are inversely proportional to the distance r from the material particle, they are not defined on the material particle but only on the surrounding spacetime, while they possess intrinsic angular momentum (spin), given respectively by

$$-i\hbar\frac{\partial\Phi}{\partial\delta}\boldsymbol{\gamma}$$

$$\frac{i\hbar}{\sin\delta}\frac{\partial\Phi}{\partial\omega}\boldsymbol{\beta}.$$

The total intrinsic angular momentum $\mathbf{S}$ of the generalized photon is given by relation

$$\mathbf{S}=\frac{i\hbar}{\sin\delta}\frac{\partial\Phi}{\partial\omega}\boldsymbol{\beta}-i\hbar\frac{\partial\Phi}{\partial\delta}\boldsymbol{\gamma}.$$

The intrinsic angular momentum of the generalized photon exhibits some remarkable properties. The first is that it does not depend on the distance r from the material particle, while it is also defined on the material particle itself. Furthermore, the component

$$S_u=i\hbar\frac{\partial\Phi}{\partial\omega}$$

in the direction of the velocity of the material particle, remains invariant under the action of the Lorentz-Einstein transformations and is, therefore, constant in all inertial reference frames. Another property of the intrinsic angular momentum of the generalized photon is that it does not vanish even if we consider that the material particle is motionless. In other words, the generalized photon carries intrinsic angular momentum even in the inertial reference frame in which the material particle is at rest. In that sense, we can characterize the intrinsic angular momentum of the generalized photon as "rest angular momentum". One final property, which is not included in the present edition is the following: during the interaction of the generalized photon with a material particle, the variation $\boldsymbol{\Delta S}$ of the angular momentum of the generalized photon manifests a component along the direction of the vector $\frac{\boldsymbol{v}}{c}$.

Of particular interest is the fact that the generalized photon, in its general version, implies the existence of rest energy in the surrounding spacetime of the material particle. The existence of this energy results as a general consequence of the equations of the theory of selfvariations.

We remind that the law of the selfvariations has been stated on the basis of the total energy E_s and the total momentum $\mathbf{P_s}$ of the generalized photons emitted simultaneously and in all directions by the material particle. We can easily prove that between the energy E_s and the momentum $\mathbf{P_s}$ the following relation holds

$$\mathbf{P_s}=E_s\frac{\mathbf{u}}{c^2}$$

where $\mathbf{u}=\mathbf{u}(w)$ is the velocity of the material particle at the moment of emission of the generalized photons. The energy E_s is always correlated with a rest energy E_0 ($E_0\neq 0$) through equation $E_s=\gamma E_0$, where $\gamma=\frac{1}{\sqrt{1-\frac{u^2}{c^2}}}$. Therefore, in the energy E_s, which results from the aggregation of the generalized photons, a rest mass of $\frac{E_0}{c^2}\neq 0$ is implicit. The law of selfvariations expresses exactly the interaction between the rest mass m_0 of the material particle, and the rest mass $\frac{E_0}{c^2}$ that results from the aggregation of the generalized photons.

The physical object that results from the aggregation of the generalized photons, always accompanies the material particle. Because of this, we named it "accompanying particle". The accompanying particle has rest mass $\frac{E_0}{c^2}$, while in the part of spacetime it occupies it holds that $dS^2=0$. The combination $\frac{E_0}{c^2}\neq 0$ and $dS^2=0$, leads to the conclusion that the accompanying particle corresponds to an intermediate state between "matter" ($\frac{E_0}{c^2}\neq 0$) and the "photon" ($dS^2=0$). This intermediate state of matter is the cause of quantum phenomena, and its prediction constitutes one of the most important results of the theory of selfvariations.

In Nature, the system material particle-accompanying particle exists and behaves as a "generalized particle" which extends in a part of spacetime. The part of space occupied by the generalized particle can be the point where the material particle is located, or it can extend up to an infinite distance away from the material particle. In the part of spacetime where the generalized particle extends, the trajectories and velocities of the generalized photons are altered with respect to the strictly defined trajectories and velocities presented in Figure 1. There is an extreme case where the concepts of trajectory and velocity of the generalized photon become meaningless; they are not defined. The same is true for the trajectory and velocity of the material particle in case it is located in the part of spacetime occupied by the generalized particle. This prediction provides us with the basic idea about the method we have to develop in order to study the generalized particle.

One way in which to study the internal structure and physical properties of the generalized particle, is to eliminate the velocity, which also represents the trajectory, from the equa-

tions of the theory of selfvariations. This elimination of the velocity can be accomplished in several ways. One is to introduce into the equations of the theory of selfvariations the potential energy U of the material particle. The resulting equation is the time-independent wave equation of Schrödinger

$$\nabla^2\Psi = -\frac{2m_0(\varepsilon - U)}{\hbar^2}\Psi.$$

The differential equations of the theory of selfvariations are of first order. When we convert them to second order equations, we can eliminate the velocity without having to introduce potential energy, or any other physical quantity, into the equations. The elimination of velocity leads to the Klein-Gordon equation. As a special case of the Klein-Gordon for $m_0 = 0$, we get the wave equation

$$\nabla^2\Psi - \frac{\partial^2\Psi}{c^2\partial t^2} = 0$$

which appears in Maxwell's theory of electromagnetism.

Observing the way in which we use Schrödinger's operators in quantum mechanics, we realize that, what we are primarily doing, is to eliminate the kinematic characteristics of the material particle from the resulting differential equations. Dirac does the same thing in the method he develops, in combination, of course, with his additional assumptions, in order to derive his eponymous equation.

In order to study the internal structure of the generalized particle we have to answer specific questions. These questions, and more generally all the issues concerning the generalized particle, are completely different from the ones we usually have to answer when we study physical reality.

The material particle can be located at any position in the part of spacetime it occupies. Judging by the success of quantum mechanics and by the high accuracy calculations it permits, we conclude that statistical interpretation is one way of studying the internal structure of the generalized particle. However, the theory of selfvariations poses a question, the answer to which, leads us to an unknown territory of physical reality.

In order to study the internal structure of the generalized particle we have to answer the question, how is the total rest mass of the generalized particle distributed between the material particle (m_0) and the accompanying particle $\left(\frac{E_0}{c^2}\right)$. During the quantitative determination of this particular distribution, the Schrödinger and Klein-Gordon equations show up, together with the wave equation of Maxwell's electromagnetic theory. In the part of spacetime occupied by the generalized particle, an external cause suffices to shift the rest mass towards either the material particle or the accompanying particle. In the first case, the generalized particle behaves as a material particle, which moves on a defined trajectory, with defined velocity, energy, etc. In the second case, the generalized particle spreads in spacetime, while the consequences of the aggregation of the generalized photons are intensified. This is the phenomenon we observe in the double-slit experiment.

The law of selfvariations results in the differential equation

$$\left(m_0c^2 + i\hbar\frac{\dot{m}_0}{m_0}\right)^{\bullet} = 0$$

the only unknown being the rest mass m_0 of the material particles. This simple equation contains as information and rationalizes, the totality of the cosmological data within a Universe that is flat and static, with the exception of a very slight variation of the fine structure constant predicted by the equations of the theory of selfvariations for observations at distances greater than 6×10^9 ly. The redshift z of a distant astronomical object located at distance r is given by equation

$$z = \frac{1 - A\exp\left(-\frac{kr}{c}\right)}{1 - A} - 1$$

where k is a constant and A is a scalar parameter that obeys the inequality

$$\frac{z}{1+z} < A < 1$$

for every value of the redshift z. Therefore, the value of parameter A is close to 1, with $A < 1$. The distance $r = r(z)$ of a distant astronomical object as a function of the redshift z, is given by equation

$$r = \frac{c}{k}\ln\left(\frac{A}{A - z(1-A)}\right).$$

In figure 2 we present the plot of the function $r = r(z)$ for $A = 0.900, A = 0.950, A = 0.990, A = 0.999$ up to $z = 5$. We observe that, as we increase the value of parameter A, the curve tends to become a straight line. This result is not accidental. It is proven that, for $A \to 1^-$, function $r = r(z)$ gives Hubble's law.

The energy $E(z)$ which fuels the radiance of astronomical objects, and which originates from the process of fusion, and generally from the conversion of mass into energy, is smaller than the corresponding energy E in our galaxy, according to equation

$$E(z) = \frac{E}{1+z}.$$

Therefore, the intrinsic luminosity of the astronomical object is lower than the standard luminosity we use. As a consequence, the luminosity distance R we measure is in fact greater than the real distance r of distant astronomical objects. The relevant calculations lead to equation

$$R = r\sqrt{1+z}.$$

Consideration the arithmetic values of the parameters that factor into function $R = R(z)$, we obtain equation

$$R = 5000\, z\sqrt{1+z}$$

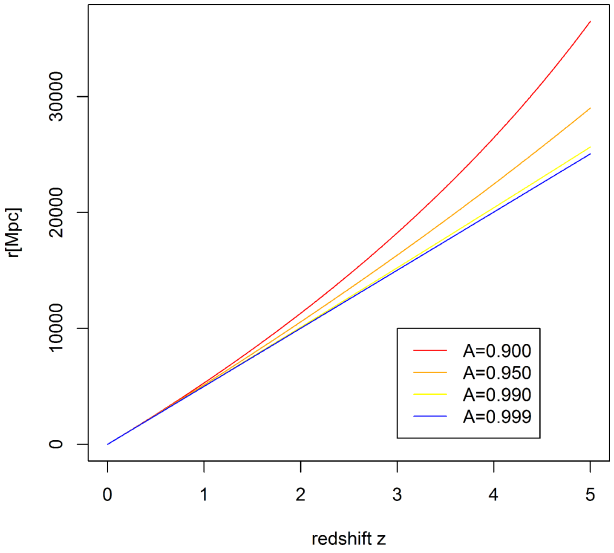

Fig. 2: The plot $r = r(z)$ of the distance of an astronomical object as a function of redshift z, for $A = 0.900, A = 0.950, A = 0.990, A = 0.999$. As the value of the parameter A is increased, the curve $r = r(z)$ tends to a straight line.

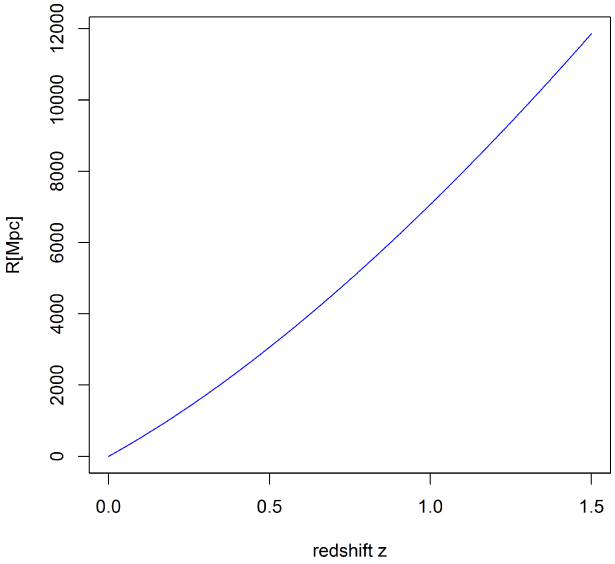

Fig. 3: The plot of the luminosity distance R of astronomical objects as a function of the redshift z. The measurement of the luminosity distances of type I_α supernova, confirms the theoretical prediction of the law of selfvariations.

where the luminosity distance R is given in Mpc. In figure 3 we present the plot of function $R = R(z)$ up to $z = 1.5$.

Type I_α supernovae are cosmological objects for which we can measure the luminosity distance at great distances. At the end of the last century, these measurements were performed by the independent scientific groups of Adam J. Riess and Saul Perlmutter. The graph that results from those measurements, exactly matches diagram in figure 3, which is theoretically predicted by the law of selfvariations. The concept of dark energy was invented in order to justify the inconsistency between the Standard Cosmological Model and diagram in figure 3.

At cosmological scales, the rest mass $m_0(r)$ with which an astronomical object exerts gravitational action at distance r from itself, is given by equation

$$m_0(r) = m_0 \frac{0.001}{1 - 0.999e^{-2\times10^{-7}r}}$$

where m_0 is the laboratory value of the rest mass. The distance r is measured in Mpc.

For values of r of the order of kpc, it turns out that $m_0 = m_0(r)$. For $r = 100$ kpc we get $m_0(r) = 0.99999\ m_0$. Consequently, the strength of the gravitational interaction is not affected on the scale of galactic distances. The selfvariations do not affect the stability of the solar system and of galaxies.

On the contrary, at distances of the order of magnitude of Mpc, a clearly smaller value of mass $m_0(r)$ compared to m_0, is predicted. For $r = 100$ Mpc we get $m_0(r) = 0.98\ m_0$. For even larger distances, the ratio $\frac{m_0(r)}{m_0}$ becomes even smaller. For an astronomical object located at a distance corresponding to redshift $z = 9$, it is $\frac{m_0(r)}{m_0} = 0.1$. The strength of the gravitational interaction exerted by an astronomical object with $z = 9$ on our galaxy is just 10% of the expected. For still greater distances, the gravitational interaction practically vanishes. This is why gravity cannot play the role attributed to it by the Standard Cosmological Model.

The Thomson scattering coefficient

$$\sigma_\tau = \frac{8\pi}{3}\frac{q^4}{m_0^2c^2}$$

as well as the Klein-Nishina scattering coefficient

$$\sigma = \frac{3}{8}\sigma_\tau \frac{m_0c^2}{E}\left[\ln\left(\frac{2E}{m_0c^2}\right) + \frac{1}{2}\right]$$

obtain different values, namely

$$\sigma_\tau(r) = \frac{8\pi}{3}\frac{q^4(r)}{m_0^2(r)c^2}$$

and

$$\sigma(r) = \frac{3}{8}\sigma_\tau(r)\frac{m_0(r)c^2}{E(r)}\left[\ln\left(\frac{2E(r)}{m_0(r)c^2}\right) + \frac{1}{2}\right]$$

respectively, at distant astronomical objects. The mathematical calculations give

$$\frac{\sigma_\tau(r)}{\sigma_\tau} = \frac{\sigma_\tau}{\sigma} = \left(\frac{1 - A\exp\left(-\frac{kr}{c}\right)}{1 - A}\right)^2.$$

At very large distances ($r \to \infty$), and equivalently for the very early Universe, we get

$$\frac{\sigma_\tau(r \to \infty)}{\sigma_\tau} = \frac{\sigma(r \to \infty)}{\sigma} = \left(\frac{1}{1-A}\right)^2.$$

Because of the inequality $\frac{z}{1+z} < A < 1$ we see that $A \to 1^-$ and, therefore, the Thomson and Klein-Nishina scattering coefficients obtain enormous values in the very early Universe. Consequently, in its very early stages, the Universe went through a phase during which it was opaque to electromagnetic radiation. The cosmic microwave background radiation originates from that period. The theory of selfvariations predicts that, in that phase, the temperature of the Universe was slightly above 0 K. Furthermore, it predicts that the cosmic microwave background radiation originates from the whole extent of the space occupied by the Universe.

The ionization and excitation energy $X_n(r) = X_n(z)$ of the atoms of distant astronomical objects differs from the laboratory value X_n according to equation

$$X_n(z) = \frac{X_n}{1+z}.$$

This equation has consequences regarding the degree of ionization of distant astronomical objects. In other words, the redshift z affects the degree of ionization of atoms in distant astronomical objects. Boltzmann's formula

$$\frac{N_n}{N_1} = \frac{g_n}{g_1} \exp\left(-\frac{X_n}{KT}\right)$$

gives the number of excited atoms N_n, that occupy the energy level n on a stellar surface which is in thermodynamic equilibrium. With X_n we denote the excitation energy from the ground energy level 1 to the energy level n, T denotes the temperature of the stellar surface in Kelvins $K = 1.38 \times 10^{-23} \frac{J}{K}$ is Boltzmann's constant, and g_n is the degree of degeneracy of energy level n (that is, the number of energy levels in which the energy level n splits in a magnetic field). At distant astronomical objects Boltzmann's formula becomes

$$\frac{N_n}{N_1} = \frac{g_n}{g_1} \exp\left(-\frac{X_n}{KT(1+z)}\right).$$

From this equation it follows that the degree of ionization at distant astronomical objects is greater than expected. The mathematical calculations lead to the conclusion that the Universe went through a phase of ionization. The dependence of the degree of ionization, as well as of the Thomson and Klein-Nishina scattering coefficients, on the redshift z, demands an overall re-evaluation of the electromagnetic spectra we receive from distant astronomical objects.

The law of selfvariations correctly predicts the structures in the Universe. It predicts the monstrous webs of matter in between vast expanses of empty space which we observe with current observational instruments. At smaller scales, it predicts galaxies and galactic clusters.

The theory of selfvariations also solves a fundamental problem concerning physical reality, which the physical theories of the last century were unable to solve: the arrow of time is included within the equations of the theory of selfvariations. The Universe comes from the vacuum and evolves towards a particular direction defined by the selfvariations. As mentioned earlier, at cosmological scales, all the equations resulting from the law of selfvariations give at the limit, for $r \to \infty$, that the initial form of the Universe only slightly differs from the vacuum at a temperature of 0 K. The origin of matter from the vacuum, in combination with the principles of conservation, with which the law of selfvariations agrees, necessitate that the energy content of the Universe remains zero. The selfvariations continually "remove" the Universe from the state of the vacuum, while at the same time the Universe remains consistent with its origin.

In contrast to what happens at the macrocosm, the equations predict that in the laboratory the arrow of time does not exist. This prediction definitively solves the problem with the arrow of time.

A measure of the future evolution of the Universe is the rate of increase of the redshift z predicted by the law of selfvariations. Substituting the arithmetic values of the parameters into the corresponding equation, we get

$$\dot{z} = z \cdot 6.3 \times 10^{-11} \text{year}^{-1}.$$

It is very characteristic the fact that one simple differential equation, having as a unique unknown the rest mass, contains as information, and at the same time justifies, the totality of the cosmological data, as we observe and record them, from the time of Hubble up to the present. Generally, the equations of the theory of Selfvariations contain an extremely large amount of data and information.

2 The study of the selfvariations for an arbitrarily moving point particle

2.1 Introduction

In this article we present the fundamental study for the mathematical background of the theory of selfvariations. We prove a set of equations which permits us the following: We can represent in the surrounding spacetime of a material particle any kinematic characteristic which concerns the material particle. At every point of spacetime, the velocity, the acceleration, the tangent vector, the curvature and the torsion of the trajectory of the material particle can be mapped in a one-to-one correspondence. This mapping allows us to take the next step: we exactly determine the contribution of the material particle to the energy content of the surrounding spacetime. What emerges is a continuous interaction of every material particle with the surrounding spacetime. The equations are proven for a material point particle in arbitrary motion. We

present a more general statement of the equations in the paragraph 8.

2.2 Arbitrarily moving material point particle

The theory of selfvariations is based upon two hypotheses which are taken as axioms.

a) The rest mass and the electric charge of the material particles increase slightly with the passage of time. We shall call this increase "selfvariations".

b) The consequences of the selfvariations propagate within the four-dimensional spacetime with a vanishing four-dimensional arc length:

$$dS^2 = 0.$$

In an inertial frame of reference, according to the second postulate, the velocity of propagation of the selfvariations remains constant as a vector

$$\boldsymbol{\upsilon} = \begin{bmatrix} \upsilon_x \\ \upsilon_y \\ \upsilon_z \end{bmatrix} = \text{constant}. \tag{1}$$

This vector has magnitude

$$\|\boldsymbol{\upsilon}\| = \sqrt{\upsilon_x^2 + \upsilon_y^2 + \upsilon_z^2} = c. \tag{2}$$

The selfvariations cause energy changes to every material particle and, as a consequence, energy, linear momentum and angular momentum propagate into the surrounding spacetime.

We shall later call the carrier of this energy, "generalized photon". Initially, we will refer to the generalized photon as a signal emitted by the material particle, moving with velocity $\boldsymbol{\upsilon}$, and, as our study advances, its properties as a real physical object will be revealed.

We consider an inertial frame of reference $S(0, x, y, z, t)$ and a material point particle moving with velocity $\mathbf{u}$ as depicted in figure 4.

At moment t, when the particle is located at point

$$P(x_p(t), y_p(t), z_p(t), t),$$

the rest mass m_0 and the electric charge q of the particle act at point $A(x, y, z, t)$ with the value they had at time $\Delta t = \frac{\|\mathbf{r}\|}{c} = \frac{r}{c}$, when the material particle was located at $E(x_p(t-\frac{r}{c}), y_p(t-\frac{r}{c}), z_p(t-\frac{r}{c}), t-\frac{r}{c})$. During the time interval $\Delta t = \frac{r}{c}$ the material particle moved from point E to point P, while the generalized photon moved from point E to point A. We now denote

$$w = t - \frac{r}{c}. \tag{3}$$

Hence, the coordinates of E are

$$E(x_p(w), y_p(w), z_p(w), w). \tag{4}$$

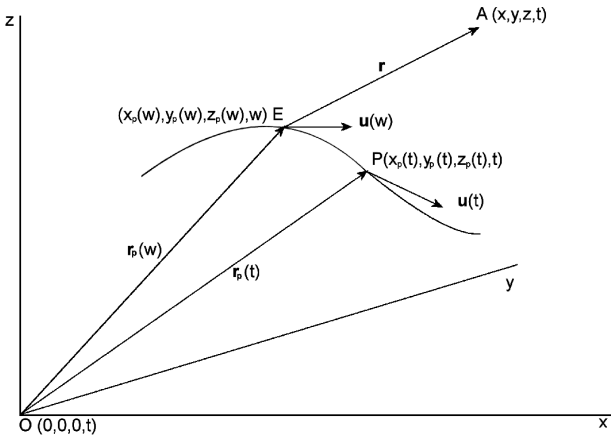

Fig. 4: Material point particle in arbitrary motion. As the material particle moves from point $E(x_p(w), y_p(w), z_p(w), w)$ to point $P(x_p(t), y_p(t), z_p(t), t)$, a generalized photon moves from point $E(x_p(w), y_p(w), z_p(w), w)$ to point $A(x, y, z, t)$.

The vector $\mathbf{r} = \overrightarrow{EA}$ of figure 4 is given by

$$\mathbf{r} = \overrightarrow{EA} = \begin{bmatrix} x - x_p(w) \\ y - y_p(w) \\ z - z_p(w) \end{bmatrix}. \tag{5}$$

The velocity of propagation of the selfvariations $\boldsymbol{\upsilon}$ is given by

$$\boldsymbol{\upsilon} = \frac{c}{r}\mathbf{r} = \frac{c}{r}\begin{bmatrix} x - x_p(w) \\ y - y_p(w) \\ z - z_p(w) \end{bmatrix}. \tag{6}$$

Here,

$$r = \|\mathbf{r}\| = \sqrt{\big(x - x_p(w)\big)^2 + \big(y - y_p(w)\big)^2 + \big(z - z_p(w)\big)^2}. \tag{7}$$

The velocity $\mathbf{u} = \mathbf{u}(w)$ of the material particle at point E, where it emitted the generalized photon, is

$$\mathbf{u} = \mathbf{u}(w) = \begin{bmatrix} \dfrac{dx_p(w)}{dw} \\ \dfrac{dy_p(w)}{dw} \\ \dfrac{dz_p(w)}{dw} \end{bmatrix}. \tag{8}$$

From equation (7) we have

$$\begin{aligned} \frac{\partial r}{\partial t} &= \frac{1}{2r}\left[2\left((x - x_p(w))\left(-\frac{dx_p(w)}{dw}\frac{\partial w}{\partial t}\right)\right)\right] \\ &+ \frac{1}{2r}\left[2\left((y - y_p(w))\left(-\frac{dy_p(w)}{dw}\frac{\partial w}{\partial t}\right)\right)\right] \\ &+ \frac{1}{2r}\left[2\left((z - z_p(w))\left(-\frac{dz_p(w)}{dw}\frac{\partial w}{\partial t}\right)\right)\right]. \end{aligned}$$

Taking into account equations (5) and (6) we have

$$\frac{\partial r}{\partial t} = -\frac{1}{r}(\mathbf{r}\cdot\mathbf{u})\frac{\partial w}{\partial t}.$$

And with equation (3) we get

$$\frac{\partial r}{\partial t} = -\frac{1}{r}(\mathbf{r}\cdot\mathbf{u})\left(1-\frac{\partial r}{c\partial t}\right).$$

Taking into consideration that $\frac{\mathbf{r}}{r} = \frac{\boldsymbol{\upsilon}}{c}$, as deduced by equation (6) we obtain

$$\frac{\partial r}{\partial t} = -\frac{\boldsymbol{\upsilon}\cdot\mathbf{u}}{c}\left(1-\frac{\partial r}{c\partial t}\right)$$

and finally

$$\frac{\partial r}{\partial t} = -\frac{\boldsymbol{\upsilon}\cdot\mathbf{u}}{c\left(1-\dfrac{\boldsymbol{\upsilon}\cdot\mathbf{u}}{c^2}\right)} \tag{9}$$

where $\mathbf{u} = \mathbf{u}(w)$ and $\boldsymbol{\upsilon}\cdot\mathbf{u} = u_x u_x + u_y u_y + u_z u_z$.

Similarly, starting from equation (7) and differentiating with respect to x, y, z we get

$$\nabla r = \begin{bmatrix} \dfrac{\partial r}{\partial x} \\ \dfrac{\partial r}{\partial y} \\ \dfrac{\partial r}{\partial z} \end{bmatrix} = \frac{1}{1-\dfrac{\boldsymbol{\upsilon}\cdot\mathbf{u}}{c^2}}\frac{\boldsymbol{\upsilon}}{c}. \tag{10}$$

From equation (3) we obtain initially

$$\frac{\partial w}{\partial t} = \frac{1}{1-\dfrac{\boldsymbol{\upsilon}\cdot\mathbf{u}}{c^2}}. \tag{11}$$

Similarly, from equation (3) we have

$$\nabla w = \nabla\left(t-\frac{r}{c}\right) = -\frac{1}{c}\nabla r$$

and, in combination with equation (10), we get

$$\nabla w = -\frac{1}{c^2\left(1-\dfrac{\boldsymbol{\upsilon}\cdot\mathbf{u}}{c^2}\right)}\boldsymbol{\upsilon}. \tag{12}$$

From equation (7) and after differentiating with respect to x, we get

$$\begin{aligned}\frac{\partial r}{\partial x} &= \frac{1}{2r}\left[2(x-x_p(w))\left(1-\frac{\partial x_p(w)}{\partial x}\right)\right]- \\ &-\frac{1}{2r}\left[2(y-y_p(w))\frac{\partial y_p(w)}{\partial x}\right]- \\ &-\frac{1}{2r}\left[2(z-z_p(w))\frac{\partial z_p(w)}{\partial x}\right].\end{aligned}$$

Equivalently,

$$\begin{aligned}\frac{\partial r}{\partial x} &= \frac{1}{r}\left[(x-x_p(w))\left(1-\frac{dx_p(w)}{dw}\frac{\partial w}{\partial x}\right)\right]- \\ &-\frac{1}{r}\left[(y-y_p(w))\left(\frac{dy_p(w)}{dw}\frac{\partial w}{\partial x}\right)\right]- \\ &-\frac{1}{r}\left[(z-z_p(w))\left(\frac{dz_p(w)}{dw}\frac{\partial w}{\partial x}\right)\right]\end{aligned}$$

and also,

$$\begin{aligned}\frac{\partial r}{\partial x} &= \frac{x-x_p(w)}{r}-\frac{1}{r}\frac{\partial w}{\partial x}\left[(x-x_p(w))\left(\frac{dx_p(w)}{dw}\right)\right]- \\ &-\frac{1}{r}\frac{\partial w}{\partial x}\left[(y-y_p(w))\left(\frac{dy_p(w)}{dw}\right)+(z-z_p(w))\left(\frac{dz_p(w)}{dw}\right)\right].\end{aligned}$$

Taking into account equations (8) and (6) we arrive at

$$\frac{\partial r}{\partial x} = \frac{\upsilon_x}{c}-\frac{\boldsymbol{\upsilon}\cdot\mathbf{u}}{c}\frac{\partial w}{\partial x}$$

and substituting

$$\frac{\partial w}{\partial x} = -\frac{\upsilon_x}{c^2\left(1-\dfrac{\boldsymbol{\upsilon}\cdot\mathbf{u}}{c^2}\right)^2}$$

as inferred from equation (12), we finally obtain

$$\frac{\partial r}{\partial x} = -\frac{1}{c\left(1-\dfrac{\boldsymbol{\upsilon}\cdot\mathbf{u}}{c^2}\right)}\upsilon_x. \tag{13}$$

Following the same procedure differentiating with respect to y and z, we finally have

$$\nabla r = \frac{1}{\left(1-\dfrac{\boldsymbol{\upsilon}\cdot\mathbf{u}}{c^2}\right)}\frac{\boldsymbol{\upsilon}}{c}. \tag{14}$$

Differentiating with respect to time t, we obtain from equation (5)

$$\frac{\partial\mathbf{r}}{\partial t} = \begin{bmatrix} -\dfrac{\partial x_p(w)}{\partial t} \\ -\dfrac{\partial y_p(w)}{\partial t} \\ -\dfrac{\partial x_p(w)}{\partial t} \end{bmatrix} = \begin{bmatrix} -\dfrac{dx_p(w)}{dw}\dfrac{\partial w}{\partial t} \\ -\dfrac{dy_p(w)}{dw}\dfrac{\partial w}{\partial t} \\ -\dfrac{dx_p(w)}{dw}\dfrac{\partial w}{\partial t} \end{bmatrix}.$$

Taking into consideration equation (8) $\frac{\partial\mathbf{r}}{\partial t} = -\frac{\partial w}{\partial t}\mathbf{u}$, and in combination with equation (11), we finally get

$$\frac{\partial\mathbf{r}}{\partial t} = -\frac{1}{\left(1-\dfrac{\boldsymbol{\upsilon}\cdot\mathbf{u}}{c^2}\right)}\mathbf{u}. \tag{15}$$

From equation (6) we successively obtain

$$\boldsymbol{\upsilon} = \frac{c}{r}\mathbf{r}$$

$$\frac{\partial \boldsymbol{\upsilon}}{\partial t} = -\frac{c}{r^2}\frac{\partial r}{\partial t}\mathbf{r} + \frac{c}{r}\frac{\partial \mathbf{r}}{\partial t}$$

$$\frac{\partial \boldsymbol{\upsilon}}{\partial t} = -\frac{1}{r}\frac{\partial r}{\partial t}\boldsymbol{\upsilon} + \frac{c}{r}\frac{\partial \mathbf{r}}{\partial t} \tag{16}$$

taking into account that $\boldsymbol{\upsilon} = \frac{c}{r}\mathbf{r}$. Substituting into equation (16) the quantity $\frac{\partial r}{\partial t}$, from equation (9), and $\frac{\partial \mathbf{r}}{\partial t}$, from (15), we finally obtain relation

$$\frac{\partial \boldsymbol{\upsilon}}{\partial t} = \frac{c}{r\left(1 - \dfrac{\boldsymbol{\upsilon}\cdot\mathbf{u}}{c^2}\right)}\left[\frac{(\boldsymbol{\upsilon}\cdot\mathbf{u})}{c^2}\boldsymbol{\upsilon} - \mathbf{u}\right]. \tag{17}$$

Starting from equation (6) we get

$$\upsilon_x = \frac{c}{r}(x - x_p(w))$$

and differentiating with respect to x we get

$$\frac{\partial \upsilon_x}{\partial x} = -\frac{c}{r^2}\frac{\partial r}{\partial x}(x - x_p(w)) + \frac{c}{r}\left(1 - \frac{\partial x_p(w)}{\partial x}\right)$$

$$\frac{\partial \upsilon_x}{\partial x} = -\frac{c}{r^2}\frac{\partial r}{\partial x}(x - x_p(w)) + \frac{c}{r}\left(1 - \frac{dx_p(w)}{dw}\frac{\partial w}{\partial x}\right).$$

Since $\frac{dx_p(w)}{dw} = u_x$, as arises from equation (8), we have that

$$\frac{\partial \upsilon_x}{\partial x} = -\frac{c}{r^2}\frac{\partial r}{\partial x}(x - x_p(w)) + \frac{c}{r}\left(1 - u_x\frac{\partial w}{\partial x}\right)$$

and considering that $\frac{\partial r}{\partial x} = -\frac{1}{c\left(1-\frac{\boldsymbol{\upsilon}\cdot\mathbf{u}}{c^2}\right)}\upsilon_x$ from equation (13), and that $\frac{\partial w}{\partial x} = \frac{1}{c^2\left(1-\frac{\boldsymbol{\upsilon}\cdot\mathbf{u}}{c^2}\right)}\upsilon_x$ from equation (12), we get

$$\frac{\partial \upsilon_x}{\partial x} = -\frac{\upsilon_x^2}{cr\left(1 - \dfrac{\boldsymbol{\upsilon}\cdot\mathbf{u}}{c^2}\right)} + \frac{c}{r}\left(1 + \frac{\upsilon_x u_x}{c^2\left(1 - \dfrac{\boldsymbol{\upsilon}\cdot\mathbf{u}}{c^2}\right)}\right)$$

and finally

$$\frac{\partial \upsilon_x}{\partial x} = \frac{c}{r} + \frac{\upsilon_x(u_x - \upsilon_x)}{cr\left(1 - \dfrac{\boldsymbol{\upsilon}\cdot\mathbf{u}}{c^2}\right)}. \tag{18}$$

Working similarly, we finally obtain

$$\frac{\partial \upsilon_i}{\partial x_j} = \begin{cases} \dfrac{c}{r} + \dfrac{\upsilon_i(u_i - \upsilon_i)}{cr\left(1 - \dfrac{\boldsymbol{\upsilon}\cdot\mathbf{u}}{c^2}\right)}, & \text{for } i = j \\ \dfrac{\upsilon_j(u_i - \upsilon_i)}{cr\left(1 - \dfrac{\boldsymbol{\upsilon}\cdot\mathbf{u}}{c^2}\right)}, & \text{for } i \neq j \end{cases} \tag{19}$$

where $i, j = 1, 2, 3$ and $(x_1, x_2, x_3) = (x, y, z)$.

Equations (19) can be summarized in equation [1–3]

$$\operatorname{grad}\boldsymbol{\upsilon} = \frac{\mathrm{c}}{\mathrm{r}}\mathrm{I} + \frac{1}{\mathrm{r}\left(1 - \dfrac{\boldsymbol{\upsilon}\cdot\mathbf{u}}{\mathrm{c}^2}\right)\mathrm{c}} \otimes (\mathbf{u} - \boldsymbol{\upsilon}) \tag{20}$$

where,

$$\operatorname{grad}\boldsymbol{\upsilon} = \begin{bmatrix} \dfrac{\partial \upsilon_x}{\partial x} & \dfrac{\partial \upsilon_x}{\partial y} & \dfrac{\partial \upsilon_x}{\partial z} \\ \dfrac{\partial \upsilon_y}{\partial x} & \dfrac{\partial \upsilon_y}{\partial y} & \dfrac{\partial \upsilon_y}{\partial z} \\ \dfrac{\partial \upsilon_z}{\partial x} & \dfrac{\partial \upsilon_z}{\partial y} & \dfrac{\partial \upsilon_z}{\partial z} \end{bmatrix}$$

$$I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\boldsymbol{\alpha}\otimes\boldsymbol{\beta} = \begin{bmatrix} \alpha_1\beta_1 & \alpha_2\beta_1 & \alpha_3\beta_1 \\ \alpha_1\beta_2 & \alpha_2\beta_2 & \alpha_3\beta_2 \\ \alpha_1\beta_3 & \alpha_2\beta_3 & \alpha_3\beta_3 \end{bmatrix}. \tag{21}$$

This holds for any two arbitrary vectors

$$\boldsymbol{\alpha} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} \quad \text{and} \quad \boldsymbol{\beta} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix}.$$

We now have $\nabla\boldsymbol{\upsilon} = \frac{\partial \upsilon_x}{\partial x} + \frac{\partial \upsilon_y}{\partial y} + \frac{\partial \upsilon_z}{\partial z}$, and from equations (19) we get

$$\nabla\boldsymbol{\upsilon} = \frac{3c}{r} + \frac{\upsilon_x(u_x - \upsilon_x) + \upsilon_y(u_y - \upsilon_y) + \upsilon_z(u_z - \upsilon_z)}{cr\left(1 - \dfrac{\boldsymbol{\upsilon}\cdot\mathbf{u}}{c^2}\right)}$$

$$\nabla\boldsymbol{\upsilon} = \frac{3c}{r} + \frac{\upsilon_x u_x + \upsilon_y u_y + \upsilon_z u_z - \left(\upsilon_x^2 + \upsilon_y^2 + \upsilon_z^2\right)}{cr\left(1 - \dfrac{\boldsymbol{\upsilon}\cdot\mathbf{u}}{c^2}\right)}$$

and since $\upsilon_x^2 + \upsilon_y^2 + \upsilon_z^2 = c^2$ and $\upsilon_x u_x + \upsilon_y u_y + \upsilon_z u_z = \boldsymbol{\upsilon}\cdot\mathbf{u}$, we see that

$$\nabla\boldsymbol{\upsilon} = \frac{3c}{r} + \frac{\boldsymbol{\upsilon}\cdot\mathbf{u} - c^2}{cr\left(1 - \dfrac{\boldsymbol{\upsilon}\cdot\mathbf{u}}{c^2}\right)}.$$

Finally, we arrive at relation

$$\nabla\boldsymbol{\upsilon} = \frac{2c}{r}. \tag{22}$$

Now, we consider the curl of vector $\boldsymbol{\upsilon}$

$$\nabla\times\boldsymbol{\upsilon} = \operatorname{curl}\boldsymbol{\upsilon} = \begin{bmatrix} \dfrac{\partial \upsilon_z}{\partial x} - \dfrac{\partial \upsilon_y}{\partial z} \\ \dfrac{\partial \upsilon_x}{\partial z} - \dfrac{\partial \upsilon_z}{\partial x} \\ \dfrac{\partial \upsilon_y}{\partial x} - \dfrac{\partial \upsilon_x}{\partial y} \end{bmatrix}. \tag{23}$$

Taking into account equations (19) we obtain

$$\nabla\times\boldsymbol{\upsilon} = \operatorname{curl}\boldsymbol{\upsilon} = \frac{1}{\mathrm{cr}\left(1 - \dfrac{\boldsymbol{\upsilon}\cdot\mathbf{u}}{\mathrm{c}^2}\right)}(\boldsymbol{\upsilon}\times\mathbf{u}) \tag{24}$$

where,

$$\boldsymbol{\upsilon}\times\mathbf{u}=\begin{bmatrix}\upsilon_y u_z-\upsilon_z u_y\\ \upsilon_z u_x-\upsilon_x u_z\\ \upsilon_x u_y-\upsilon_y u_x\end{bmatrix}.$$

We now consider the acceleration vector

$$\boldsymbol{\alpha}=\boldsymbol{\alpha}(w)=\frac{d\mathbf{u}(w)}{dw}=\begin{bmatrix}\dfrac{d\mathbf{u_x}(w)}{dw}\\ \dfrac{d\mathbf{u_y}(w)}{dw}\\ \dfrac{d\mathbf{u_z}(w)}{dw}\end{bmatrix} \tag{25}$$

of the material particle at the moment w, located at point E of figure 4. We have that

$$\frac{\partial u_x}{\partial t}=\frac{\partial u_x(w)}{\partial t}=\frac{du_x(w)}{dw}\frac{\partial w}{\partial t}=\alpha_x\frac{\partial w}{\partial t}$$

and since, from equation (11), it is $\frac{\partial w}{\partial t}=\frac{1}{1-\frac{\upsilon\cdot\mathbf{u}}{c^2}}$, we get $\frac{\partial u_x}{\partial t}=\frac{\alpha_x}{1-\frac{\upsilon\cdot\mathbf{u}}{c^2}}$. Working similarly for the differentials $\frac{\partial u_y}{\partial t}$ and $\frac{\partial u_z}{\partial t}$, we get

$$\frac{\partial\mathbf{u}}{\partial t}=\frac{1}{1-\dfrac{\boldsymbol{\upsilon}\cdot\mathbf{u}}{c^2}}\boldsymbol{\alpha}. \tag{26}$$

For the differentiation of the velocity $\mathbf{u}=\mathbf{u}(w)$ with respect to x,y,z we initially get

$$\frac{\partial u_x}{\partial x}=\frac{\partial u_x(w)}{\partial x}=\frac{du_x(w)}{dw}\frac{\partial w}{\partial x}=\alpha_x\frac{\partial w}{\partial x}.$$

Similarly, from equation (12) we have that $\frac{\partial w}{\partial x}=-\frac{u_x}{c^2\left(1-\frac{\upsilon\cdot\mathbf{u}}{c^2}\right)}$, hence $\frac{\partial u_x}{\partial x}=-\frac{\upsilon_x\alpha_x}{c^2\left(1-\frac{\upsilon\cdot\mathbf{u}}{c^2}\right)}$.

Working similarly we finally obtain

$$\frac{\partial u_i}{\partial x_j}=-\frac{\upsilon_j\alpha_i}{c^2\left(1-\dfrac{\boldsymbol{\upsilon}\cdot\mathbf{u}}{c^2}\right)}\qquad i,j=1,2,3. \tag{27}$$

Here we use the notation $(x_1,x_2,x_3)=(x,y,z)$.

From equation (27) we obtain

$$grad\mathbf{u}=-\frac{1}{r\left(1-\dfrac{\boldsymbol{\upsilon}\cdot\mathbf{u}}{c^2}\right)}\frac{\boldsymbol{\upsilon}}{c}\otimes\mathbf{u}. \tag{28}$$

We now consider the vector

$$\mathbf{b}=\mathbf{b}(w)=\frac{d\boldsymbol{\alpha}(w)}{dw}. \tag{29}$$

Working as we did in order to prove equations (17), (26) and (27), we arrive at relations

$$\frac{\partial\boldsymbol{\alpha}}{\partial t}=\frac{1}{1-\dfrac{\boldsymbol{\upsilon}\cdot\mathbf{u}}{c^2}}\mathbf{b} \tag{30}$$

$$\frac{\partial\alpha_i}{\partial x_j}=-\frac{\upsilon_j b_i}{c^2\left(1-\dfrac{\boldsymbol{\upsilon}\cdot\boldsymbol{u}}{c^2}\right)}\qquad i,j=1,2,3 \tag{31}$$

where $(x_1,x_2,x_3)=(x,y,z)$, and

$$\text{grad}\,\boldsymbol{\alpha}=-\frac{1}{\mathrm{c}\left(1-\dfrac{\boldsymbol{\upsilon}\cdot\mathbf{u}}{\mathrm{c}^2}\right)}\frac{\boldsymbol{\upsilon}}{\mathrm{c}}\otimes\mathbf{b}. \tag{32}$$

The equations of this paragraph express the fact that in every inertial reference frame the velocity $\boldsymbol{\upsilon}$ of the selfvariations remains constant as a vector with magnitude $\|\boldsymbol{\upsilon}\|=c$. It can easily be proven that all the equations are consistent with the Lorentz-Einstein transformations, as we pass from one inertial reference frame to another. The equations we have proven are fundamental for the theory of selfvariations. As we advance our study, we will find that they allow us to correlate any physical quantity defined on the material particle, with any physical quantity defined on the surrounding spacetime. Using the concept of information, we can correlate any information concerning the material particle with any information concerning the surrounding spacetime. Part of this information are the potential fields, while the quantum phenomena arise spontaneously.

2.3 The trigonometric form of the velocity of selfvariations

Starting from equation (2) we get $\left\|\frac{\boldsymbol{\upsilon}}{c}\right\|=1$ for every inertial reference frame. We express the unit vector $\frac{\boldsymbol{\upsilon}}{c}$ into the trigonometric form

$$\frac{\boldsymbol{\upsilon}}{c}=\begin{bmatrix}\dfrac{\upsilon_x}{c}\\ \dfrac{\upsilon_y}{c}\\ \dfrac{\upsilon_z}{c}\end{bmatrix}=\begin{bmatrix}\cos\delta\\ \sin\delta\cos\omega\\ \sin\delta\sin\omega\end{bmatrix} \tag{33}$$

where $\delta=\delta(x,y,z,t)$ and $\omega=\omega(x,y,z,t)$ are functions of the coordinates x,y,z,t in an inertial frame of reference $S(0,x,y,z,t)$.

From equation (33) we see that

$$\frac{\upsilon_x}{c}=\cos\delta=\frac{\boldsymbol{\upsilon}}{c}\boldsymbol{e}_1 \tag{34a}$$

$$\frac{\upsilon_y}{c}=\sin\delta\cos\omega=\frac{\boldsymbol{\upsilon}}{c}\boldsymbol{e}_2 \tag{34b}$$

$$\frac{\upsilon_z}{c}=\sin\delta\sin\omega=\frac{\boldsymbol{\upsilon}}{c}\boldsymbol{e}_3 \tag{34c}$$

where

$$\boldsymbol{e}_1=\hat{\boldsymbol{x}}=\begin{bmatrix}1\\0\\0\end{bmatrix},\quad \boldsymbol{e}_2=\hat{\boldsymbol{y}}=\begin{bmatrix}0\\1\\0\end{bmatrix},\quad \boldsymbol{e}_3=\hat{\boldsymbol{z}}=\begin{bmatrix}0\\0\\1\end{bmatrix}.$$

We now consider the vectors

$$\boldsymbol{\beta}=\begin{bmatrix}-\sin\delta\\ \cos\delta\cos\omega\\ \cos\delta\sin\omega\end{bmatrix} \tag{35}$$

and

$$\boldsymbol{\gamma} = \begin{bmatrix} 0 \\ -\sin\omega \\ \cos\omega \end{bmatrix}. \tag{36}$$

It is easily proven that the set of vectors $\{\frac{\boldsymbol{v}}{c}, \boldsymbol{\beta}, \boldsymbol{\gamma}\}$ form a right-handed orthonormal vector basis which is defined at every point A of figure 4. Furthermore, the following relations hold:

$$\begin{aligned}
&\frac{\partial}{\partial\delta}\left(\frac{\boldsymbol{v}}{c}\right) = \boldsymbol{\beta} \\
&\frac{\partial}{\partial\omega}\left(\frac{\boldsymbol{v}}{c}\right) = \sin\delta\boldsymbol{\gamma} \\
&\frac{\partial\boldsymbol{\beta}}{\partial\delta} = -\frac{\boldsymbol{v}}{c} \\
&\frac{\partial\boldsymbol{\beta}}{\partial\omega} = \cos\delta\boldsymbol{\gamma} \\
&\frac{\partial\boldsymbol{\gamma}}{\partial\delta} = 0 \\
&\frac{\partial\boldsymbol{\gamma}}{\partial\omega} = -\sin\delta\frac{\boldsymbol{v}}{c} - \cos\delta\boldsymbol{\beta}
\end{aligned} \tag{37}$$

Differentiating the vectors $\frac{\boldsymbol{v}}{c}, \boldsymbol{\beta}, \boldsymbol{\gamma}$ with respect to x, y, z, t we obtain the following equations:

$$\nabla\cdot\left(\frac{\boldsymbol{v}}{c}\right) = \beta\cdot\nabla\delta + \sin\delta\boldsymbol{\gamma}\cdot\nabla\omega \tag{38a}$$

$$\frac{\partial}{\partial t}\left(\frac{\boldsymbol{v}}{c}\right) = \frac{\partial\delta}{\partial t}\boldsymbol{\beta} + \sin\delta\frac{\partial\omega}{\partial t}\boldsymbol{\gamma} \tag{38b}$$

$$\nabla\times\frac{\boldsymbol{v}}{c} = \nabla\delta\times\boldsymbol{\beta} + \sin\delta\nabla\omega\otimes\boldsymbol{\gamma} \tag{38c}$$

$$\operatorname{grad}\frac{\boldsymbol{v}}{\mathrm{c}} = \nabla\delta\otimes\boldsymbol{\beta} + \sin\delta\nabla\omega\otimes\boldsymbol{\gamma} \tag{38d}$$

$$\nabla\cdot\boldsymbol{\beta} = -\frac{\boldsymbol{v}}{c}\nabla\delta + \cos\delta\boldsymbol{\gamma}\cdot\nabla\omega \tag{39a}$$

$$\frac{\partial\boldsymbol{\beta}}{\partial t} = -\frac{\partial\delta}{\partial t}\frac{\boldsymbol{v}}{c} + \cos\delta\frac{\partial\omega}{\partial t}\boldsymbol{\gamma} \tag{39b}$$

$$\nabla\times\boldsymbol{\beta} = \frac{\boldsymbol{v}}{c}\times\nabla\delta - \cos\delta\boldsymbol{\gamma}\times\nabla\omega \tag{39c}$$

$$\operatorname{grad}\boldsymbol{\beta} = -\nabla\delta\otimes\frac{\boldsymbol{v}}{\mathrm{c}} + \cos\delta\nabla\omega\otimes\boldsymbol{\gamma} \tag{39d}$$

$$\nabla\cdot\boldsymbol{\gamma} = -\sin\delta\frac{\boldsymbol{v}}{c}\nabla\omega - \cos\delta\boldsymbol{\beta}\cdot\nabla\omega \tag{40a}$$

$$\frac{\partial\boldsymbol{\gamma}}{\partial t} = -\sin\delta\frac{\partial\omega}{\partial t}\frac{\boldsymbol{v}}{c} - \cos\delta\frac{\partial\omega}{\partial t}\boldsymbol{\beta} \tag{40b}$$

$$\nabla\times\boldsymbol{\gamma} = \sin\delta\frac{\boldsymbol{v}}{c}\times\nabla\omega + \cos\delta\boldsymbol{\beta}\times\nabla\omega \tag{40c}$$

$$\operatorname{grad}\boldsymbol{\gamma} = -\sin\delta\nabla\omega\otimes\frac{\boldsymbol{v}}{\mathrm{c}} - \cos\delta\nabla\omega\otimes\boldsymbol{\beta}. \tag{40d}$$

We prove indicatively equation (38)(a). The rest of the equations are proven along similar lines. Taking into account equation (33) we get

$$\begin{aligned}
\nabla\cdot\left(\frac{\boldsymbol{v}}{c}\right) &= \frac{\partial}{\partial x}(\cos\delta) + \frac{\partial}{\partial y}(\sin\delta\cos\omega) + \frac{\partial}{\partial z}(\sin\delta\sin\omega) = \\
&-\sin\delta\frac{\partial\delta}{\partial x} + \cos\delta\frac{\partial\delta}{\partial y}\cos\omega + \cos\delta\frac{\partial\delta}{\partial z}\sin\omega \\
&+ 0 - \sin\delta\sin\omega\frac{\partial\omega}{\partial y} + \sin\delta\cos\omega\frac{\partial\omega}{\partial z}
\end{aligned}$$

and considering equations (35) and (36), as well as relations

$$\nabla\delta = \begin{bmatrix} \frac{\partial\delta}{\partial x} \\ \frac{\partial\delta}{\partial y} \\ \frac{\partial\delta}{\partial z} \end{bmatrix}, \qquad \nabla\omega = \begin{bmatrix} \frac{\partial\omega}{\partial x} \\ \frac{\partial\omega}{\partial y} \\ \frac{\partial\omega}{\partial z} \end{bmatrix}$$

we finally obtain

$$\nabla\cdot\left(\frac{\boldsymbol{v}}{c}\right) = \boldsymbol{\beta}\cdot\nabla\delta + \sin\delta\boldsymbol{\gamma}\cdot\omega.$$

We now expand the vector of velocity $\boldsymbol{u} = \boldsymbol{u}(w)$ with respect to the vector basis $\{\frac{\boldsymbol{v}}{c}, \boldsymbol{\beta}, \boldsymbol{\gamma}\}$ as

$$\boldsymbol{u} = \boldsymbol{u}(w) = u_1\frac{\boldsymbol{v}}{c} + u_2\boldsymbol{\beta} + u_3\boldsymbol{\gamma} = \left(\boldsymbol{u}\cdot\frac{\boldsymbol{v}}{c}\right)\frac{\boldsymbol{v}}{c} + (\boldsymbol{u}\cdot\boldsymbol{\beta})\boldsymbol{\beta} + (\boldsymbol{u}\cdot\boldsymbol{\gamma})\boldsymbol{\gamma}$$

and combining with equations (17) we get

$$\begin{aligned}
\frac{\partial}{\partial t}\left(\frac{\boldsymbol{v}}{c}\right) &= \frac{1}{r\left(1-\frac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}\right)}\left[\frac{(\boldsymbol{v}\cdot\boldsymbol{u})}{c}\frac{\boldsymbol{v}}{c} - \left(\boldsymbol{u}\frac{\boldsymbol{v}}{c}\right)\frac{\boldsymbol{v}}{c} - (\boldsymbol{u}\cdot\boldsymbol{\beta})\boldsymbol{\beta} - (\boldsymbol{u}\cdot\boldsymbol{\gamma})\boldsymbol{\gamma}\right] \\
\frac{\partial}{\partial t}\left(\frac{\boldsymbol{v}}{c}\right) &= \frac{1}{r\left(1-\frac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}\right)}\left[(\boldsymbol{u}\cdot\boldsymbol{\beta})\boldsymbol{\beta} + (\boldsymbol{u}\cdot\boldsymbol{\gamma})\boldsymbol{\gamma}\right].
\end{aligned}$$

Considering equations (38)(b) we get

$$\frac{\partial\delta}{\partial t}\boldsymbol{\beta} + \sin\delta\frac{\partial\omega}{\partial t}\boldsymbol{\gamma} = -\frac{1}{r\left(1-\frac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}\right)}\left[(\boldsymbol{u}\cdot\boldsymbol{\beta})\boldsymbol{\beta} + (\boldsymbol{u}\cdot\boldsymbol{\gamma})\boldsymbol{\gamma}\right]$$

and finally

$$\frac{\partial\delta}{\partial t} = -\frac{\boldsymbol{u}\cdot\boldsymbol{\beta}}{r\left(1-\frac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}\right)} \tag{41}$$

$$\sin\delta\frac{\partial\omega}{\partial t} = -\frac{\boldsymbol{u}\cdot\boldsymbol{\gamma}}{r\left(1-\frac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}\right)} \tag{42}$$

because of the linear independence of the vectors $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$.

We now write vectors $\nabla\delta$ and $\nabla\omega$ as a linear combination of vectors $\frac{\boldsymbol{v}}{c}, \boldsymbol{\beta}, \boldsymbol{\gamma}$.

$$\nabla\delta = \lambda_1\frac{\boldsymbol{v}}{c} + K\boldsymbol{\beta} + L\boldsymbol{\gamma} \tag{43}$$

$$\nabla\omega = \lambda_2\frac{\boldsymbol{\upsilon}}{c} + M\boldsymbol{\beta} + N\boldsymbol{\gamma}. \tag{44}$$

We combine equations (17) and (20), and get relation

$$\frac{\partial}{\partial t}\left(\frac{\boldsymbol{\upsilon}}{c}\right) + \left(grad\frac{\boldsymbol{\upsilon}}{c}\right)\boldsymbol{\upsilon} = \frac{1}{r\left(1-\frac{\boldsymbol{\upsilon}\cdot\boldsymbol{u}}{c^2}\right)}\left[\left(\frac{\boldsymbol{\upsilon}\cdot\boldsymbol{u}}{c^2}\right)\boldsymbol{\upsilon} - \boldsymbol{u}\right] +$$
$$+\left[\frac{1}{r}I + \frac{1}{cr\left(1-\frac{\boldsymbol{\upsilon}\cdot\boldsymbol{u}}{c^2}\right)}\frac{\boldsymbol{\upsilon}}{c}\otimes(\boldsymbol{u}-\boldsymbol{\upsilon})\right]\boldsymbol{\upsilon} =$$
$$\frac{1}{r\left(1-\frac{\boldsymbol{\upsilon}\cdot\boldsymbol{u}}{c^2}\right)}\left[\left(\frac{\boldsymbol{\upsilon}\cdot\boldsymbol{u}}{c^2}\right)\boldsymbol{\upsilon} - \boldsymbol{u}\right] +$$
$$+\frac{1}{r}\boldsymbol{\upsilon} + \frac{1}{cr\left(1-\frac{\boldsymbol{\upsilon}\cdot\boldsymbol{u}}{c^2}\right)}\left(\frac{\boldsymbol{\upsilon}}{c}\otimes(\boldsymbol{u}-\boldsymbol{\upsilon})\right)\boldsymbol{\upsilon}.$$

Using the identity

$$(\boldsymbol{\alpha}\otimes\boldsymbol{\beta})\,\boldsymbol{c} = (\boldsymbol{\alpha}\cdot\boldsymbol{c})\,\boldsymbol{\beta} \tag{45}$$

which holds for every set of vectors $\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{c}$, we see that

$$\frac{\partial}{\partial t}\left(\frac{\boldsymbol{\upsilon}}{c}\right) + \left(\mathrm{grad}\,\frac{\boldsymbol{\upsilon}}{\mathrm{c}}\right)\boldsymbol{\upsilon} =$$
$$\frac{1}{r\left(1-\frac{\boldsymbol{\upsilon}\cdot\boldsymbol{u}}{c^2}\right)}\left[\left(\frac{\boldsymbol{\upsilon}\cdot\boldsymbol{u}}{c^2}\right)\boldsymbol{\upsilon} - \boldsymbol{u}\right] + \frac{1}{r}\boldsymbol{\upsilon} + \frac{1}{r\left(1-\frac{\boldsymbol{\upsilon}\cdot\boldsymbol{u}}{c^2}\right)}(\boldsymbol{\upsilon}-\boldsymbol{u}) =$$
$$\frac{1}{r\left(1-\frac{\boldsymbol{\upsilon}\cdot\boldsymbol{u}}{c^2}\right)}\left(\frac{\boldsymbol{\upsilon}\cdot\boldsymbol{u}}{c^2}\right)\boldsymbol{\upsilon} + \frac{1}{r}\boldsymbol{\upsilon} - \frac{1}{r\left(1-\frac{\boldsymbol{\upsilon}\cdot\boldsymbol{u}}{c^2}\right)}\boldsymbol{\upsilon} =$$
$$\frac{1}{r\left(1-\frac{\boldsymbol{\upsilon}\cdot\boldsymbol{u}}{c^2}\right)}\left[\left(\frac{\boldsymbol{\upsilon}\cdot\boldsymbol{u}}{c^2}\right) + \left(1-\frac{\boldsymbol{\upsilon}\cdot\boldsymbol{u}}{c^2}\right) - 1\right]\boldsymbol{\upsilon} = \boldsymbol{0}.$$

That is,

$$\frac{\partial}{\partial t}\left(\frac{\boldsymbol{\upsilon}}{c}\right) + \left(\mathrm{grad}\,\frac{\boldsymbol{\upsilon}}{\mathrm{c}}\right)\boldsymbol{\upsilon} = \boldsymbol{0}. \tag{46}$$

Into equation (45) we replace $\frac{\partial}{\partial t}\left(\frac{\upsilon}{c}\right)$ from equation (38)(b), and grad $\frac{\upsilon}{\mathrm{c}}$ from equation (38)(d), and obtain

$$\frac{\partial\delta}{\partial t}\boldsymbol{\beta} + \sin\delta\frac{\partial\omega}{\partial t}\boldsymbol{\gamma} + (\nabla\delta\otimes\boldsymbol{\beta} + \sin\delta\nabla\omega\otimes\boldsymbol{\gamma})\,\boldsymbol{\upsilon} = \boldsymbol{0}.$$

Using the identity (45) we get

$$\frac{\partial\delta}{\partial t}\boldsymbol{\beta} + \sin\delta\frac{\partial\omega}{\partial t}\boldsymbol{\gamma} + (\boldsymbol{u}\cdot\nabla\delta)\,\boldsymbol{\beta} + \sin\delta\,(\boldsymbol{\upsilon}\cdot\nabla\omega)\,\boldsymbol{\gamma} = 0$$

and due to the linear independence of the vectors $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ we see that

$$\frac{\partial\delta}{\partial t} + \boldsymbol{\upsilon}\cdot\nabla\delta = 0 \tag{47}$$

$$\frac{\partial\omega}{\partial t} + \boldsymbol{\upsilon}\cdot\nabla\omega = 0. \tag{48}$$

Combining equations (47) and (43) we obtain

$$\frac{\partial\delta}{\partial t} + \lambda_1 = 0$$
$$\lambda_1 = -\frac{\partial\delta}{\partial t}.$$

Through equation (41) we have that

$$\lambda_1 = \frac{\boldsymbol{u}\cdot\boldsymbol{\beta}}{r\left(1-\frac{\boldsymbol{\upsilon}\cdot\boldsymbol{u}}{c^2}\right)}$$

and replacing into equation (43) we get

$$\nabla\delta = \frac{\boldsymbol{u}\cdot\boldsymbol{\beta}}{r\left(1-\frac{\boldsymbol{\upsilon}\cdot\boldsymbol{u}}{c^2}\right)}\frac{\boldsymbol{\upsilon}}{c^2} + K\boldsymbol{\beta} + L\boldsymbol{\gamma}. \tag{49}$$

Performing the corresponding combinations, we arrive at equation

$$\nabla\omega = \frac{\boldsymbol{u}\cdot\boldsymbol{\gamma}}{\sin\delta r\left(1-\frac{\boldsymbol{\upsilon}\cdot\boldsymbol{u}}{c^2}\right)}\frac{\boldsymbol{\upsilon}}{c} + M\boldsymbol{\beta} + N\boldsymbol{\gamma}. \tag{50}$$

We shall now prove that $K = \frac{1}{r}$, $L = 0$, $M = 0$, $N = \frac{1}{r\sin\delta}$, hence equations (49) and (50) obtain their final form

$$\nabla\delta = \frac{\boldsymbol{u}\cdot\boldsymbol{\beta}}{r\left(1-\frac{\boldsymbol{\upsilon}\cdot\boldsymbol{u}}{c^2}\right)}\frac{\boldsymbol{\upsilon}}{c^2} + \frac{1}{r}\boldsymbol{\beta} \tag{51}$$

$$\nabla\omega = \frac{\boldsymbol{u}\cdot\boldsymbol{\gamma}}{\sin\delta r\left(1-\frac{\boldsymbol{\upsilon}\cdot\boldsymbol{u}}{c^2}\right)}\frac{\boldsymbol{\upsilon}}{c^2} + \frac{1}{r\sin\delta}\boldsymbol{\gamma}. \tag{52}$$

We will prove that $K = \frac{1}{r}$, $L = 0$. In a similar manner we can also calculate the factors M, N. From equation (34)(a) we successively obtain

$$\cos\delta = \frac{\upsilon_x}{c}$$
$$-\sin\delta\nabla\delta = \nabla\left(\frac{\upsilon_x}{c}\right).$$

We calculate $\nabla\left(\frac{\upsilon_x}{c}\right)$ from equations (19), hence we have

$$-\sin\delta\nabla\delta = \frac{1}{r}\boldsymbol{e}_1 - \frac{\upsilon_x - u_x}{r\left(1-\frac{\boldsymbol{\upsilon}\cdot\boldsymbol{u}}{c^2}\right)}\frac{\boldsymbol{\upsilon}}{c} \tag{53}$$

where $\boldsymbol{e}_1 = \begin{bmatrix}1\\0\\0\end{bmatrix}$.

We take the inner product of equation (53) with vector $\boldsymbol{\beta}$ and obtain

$$-\sin\delta\boldsymbol{\beta}\cdot\nabla\delta = \frac{1}{r}\boldsymbol{e}_1\cdot\boldsymbol{\beta}.$$

From equation (49) we have $\boldsymbol{\beta}\cdot\nabla\delta = K$, hence we have

$$-\sin\delta K = \frac{1}{r}\boldsymbol{e}_1\cdot\boldsymbol{\beta}.$$

From equation (35) we obtain

$$\boldsymbol{e}_1\cdot\boldsymbol{\beta} = -\sin\delta.$$

Therefore,

$$-\sin\delta K = \frac{1}{r}(-\sin\delta).$$

Finally, we obtain

$$K = \frac{1}{r}.$$

We take the inner product of equation (53) with vector $\boldsymbol{\gamma}$ and obtain

$$-\sin\delta\boldsymbol{\gamma}\cdot\nabla\delta = \frac{1}{r}\boldsymbol{e}_1\cdot\boldsymbol{\gamma}.$$

From equation (49) it holds that $\boldsymbol{\gamma}\cdot\nabla\delta = L$, hence

$$-\sin\delta L = \frac{1}{r}\boldsymbol{e}_1\cdot\boldsymbol{\gamma}.$$

From equation (36) we see that $\boldsymbol{e}_1\cdot\boldsymbol{\gamma} = 0$, therefore $-\sin\delta L = 0$, and finally $L = 0$.

The equations of this paragraph promote the theory of selfvariations considerably, and their fundamental character will become obvious as our study continues. One first fundamental conclusion emerges from equations (47) and (48). The functions $\delta = \delta(x,y,z,t)$ and $\omega = \omega(x,y,z,t)$ remain invariable on the trajectory of the generalized photon. Through equations (33), (35) and (36) we conclude that the vector basis $\{\frac{\boldsymbol{v}}{c},\boldsymbol{\beta},\boldsymbol{\gamma}\}$ accompanies without change, that is remaining constant, the motion of the generalized photon. We can, of course, straightforwardly prove that

$$\begin{aligned}
&\frac{\partial}{\partial t}\left(\frac{\boldsymbol{v}}{c}\right) + \left(\operatorname{grad}\frac{\boldsymbol{v}}{\mathrm{c}}\right)\boldsymbol{v} = 0\\
&\frac{\partial\boldsymbol{\beta}}{\partial t} + (\operatorname{grad}\boldsymbol{\beta})\,\boldsymbol{v} = 0\\
&\frac{\partial\boldsymbol{\gamma}}{\partial t} + (\operatorname{grad}\boldsymbol{\gamma})\,\boldsymbol{v} = 0
\end{aligned} \tag{54}$$

by combining equations (38), (39) and (40) with equations (51) and (52).

2.4 The generalized photon as a geometric object. Representation of the trajectory of a material point particle

In the present paragraph we shall look for points A_i in the neighborhood of point $A(x,y,z,t)$ of figure 4, for which the velocity of the generalized photon is the same with the velocity at point $A(x,y,z,t)$ at the same moment t. We use the notation

$$\overrightarrow{AA_i} = d\boldsymbol{R} \tag{55}$$

and we search for points A_i, i.e. vector $d\boldsymbol{R}$, such that

$$\boldsymbol{v}(\boldsymbol{R}+d\boldsymbol{R},t) = \boldsymbol{v}(\boldsymbol{R},t). \tag{56}$$

According to equations (33), equation (56) is equivalent to the relations

$$\delta(\boldsymbol{R}+d\boldsymbol{R},t) = \delta(\boldsymbol{R},t) \tag{57}$$

and

$$\omega(\boldsymbol{R}+d\boldsymbol{R},t) = \omega(\boldsymbol{R},t). \tag{58}$$

After expanding the functions $\delta(\boldsymbol{R},t)$ and $\omega(\boldsymbol{R},t)$ in Taylor series up to the first order terms, we obtain

$$\begin{aligned}
&\delta(\boldsymbol{R}+d\boldsymbol{R},t) = \delta(\boldsymbol{R},t) + d\boldsymbol{R}\cdot\nabla\delta\\
&\omega(\boldsymbol{R}+d\boldsymbol{R},t) = \omega(\boldsymbol{R},t) + d\boldsymbol{R}\cdot\nabla\omega.
\end{aligned}$$

Through equations (57) and (58) we have that

$$d\boldsymbol{R}\cdot\nabla\delta = 0 \tag{59}$$

$$d\boldsymbol{R}\cdot\nabla\omega = 0. \tag{60}$$

Combining equations (51) and (52) we obtain

$$\begin{aligned}
&\boldsymbol{t} = \nabla\delta\times\sin\delta\nabla\omega =\\
&\frac{\boldsymbol{u}\cdot\boldsymbol{\beta}}{r^2\left(1-\frac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}\right)}\frac{\boldsymbol{v}}{c}\times\boldsymbol{\gamma} + \frac{\boldsymbol{u}\cdot\boldsymbol{\gamma}}{r^2\left(1-\frac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}\right)}\boldsymbol{\beta}\times\frac{\boldsymbol{v}}{c} + \frac{1}{r^2}\boldsymbol{\beta}\times\boldsymbol{\gamma} =\\
&-\frac{\boldsymbol{u}\cdot\boldsymbol{\beta}}{r^2\left(1-\frac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}\right)}\boldsymbol{\beta} - \frac{\boldsymbol{u}\cdot\boldsymbol{\gamma}}{r^2\left(1-\frac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}\right)}\boldsymbol{\gamma} + \frac{1}{r^2}\frac{\boldsymbol{v}}{c}
\end{aligned}$$

taking into account that the set of the vectors $\{\frac{\boldsymbol{v}}{c},\boldsymbol{\beta},\boldsymbol{\gamma}\}$ form a right-handed orthonormal vector basis. We now have

$$\begin{aligned}
\boldsymbol{t} &= \frac{1}{r^2\left(1-\frac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}\right)}\left[\left(1-\frac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}\right)\frac{\boldsymbol{v}}{c} - \left(\frac{\boldsymbol{u}}{c}\cdot\boldsymbol{\beta}\right)\boldsymbol{\beta} - \left(\frac{\boldsymbol{u}}{c}\cdot\boldsymbol{\gamma}\right)\boldsymbol{\gamma}\right]\\
\boldsymbol{t} &= \frac{1}{r^2\left(1-\frac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}\right)}\left[\frac{\boldsymbol{v}}{c} - \left(\frac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}\right)\frac{\boldsymbol{v}}{c} - \left(\frac{\boldsymbol{u}}{c}\cdot\boldsymbol{\beta}\right)\boldsymbol{\beta} - \left(\frac{\boldsymbol{u}}{c}\cdot\boldsymbol{\gamma}\right)\boldsymbol{\gamma}\right]
\end{aligned}$$

and from equation (41) we get

$$\boldsymbol{t} = \frac{1}{r^2\left(1-\frac{\boldsymbol{v}\cdot\boldsymbol{u}}{c}\right)}\left(\frac{\boldsymbol{v}}{c}-\frac{\boldsymbol{u}}{c}\right) \neq \boldsymbol{0}. \tag{61}$$

According to equations (59) and (60) the vector $d\boldsymbol{R}$ is parallel to the vector $\boldsymbol{t}\neq\boldsymbol{0}$, hence we finally arrive at relation

$$d\boldsymbol{R} \parallel \left(\frac{\boldsymbol{v}}{c}-\frac{\boldsymbol{u}}{c}\right) \tag{62}$$

Thus, we conclude that points A and A_i, at which the generalized photon moves with the same velocity $\boldsymbol{v}$, are arranged parallel to the vector $\frac{\boldsymbol{v}}{c}-\frac{\boldsymbol{u}}{c}$. This conclusion is the result of a more general theorem, which we present in the paragraph

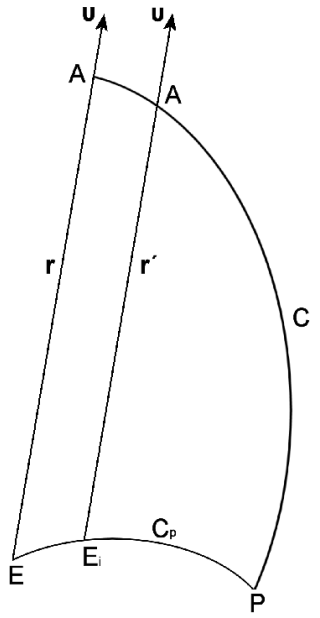

Fig. 5: A material point particle moves from point E to point P on the curved trajectory C_p in the time interval from $w = t - \frac{r}{c}$ to t. The generalized photons emitted by the material particle with the same velocity υ, in the time interval $\Delta t = t - w = \frac{r}{c}$, are on curve C at moment t.

8. For the case of a material point particle the theorem gives relation (62).

In figure 4 and for the time interval from $t - \frac{r}{c}$ to t, i.e. for $t - \frac{r}{c} \leq w \leq t$, the generalized photons emitted by the material point particle reside within a sphere with center

$$E\left(x_p\left(t-\frac{r}{c}\right), y_p\left(t-\frac{r}{c}\right), z_p\left(t-\frac{r}{c}\right), t-\frac{r}{c}\right)$$

and radius $r = \|\boldsymbol{r}\|$. During the same time interval the material particle moved from point E to point $P\left(x_p(t), y_p(t), z_p(t), t\right)$.

We now consider a point E_i in the neighborhood of point E and on the trajectory C_p of the material particle as it moves from point E to point P, from which point E_i was emitted the generalized photon which at moment t is located at point A_i, as depicted in figure 5.

Point E_i has coordinates

$$E_i\left(x_p\left(t-\frac{r'}{c}\right), y_p\left(t-\frac{r'}{c}\right), z_p\left(t-\frac{r'}{c}\right), t-\frac{r'}{c}\right),$$

where $\boldsymbol{\upsilon} = \frac{c}{r}\boldsymbol{r} = \frac{c}{r'}\boldsymbol{r}'$.

The points E, P, A appear in figure 4 as well as in figure 5, while the points E_i and A_i are shown in figure 5.

For the vector $\overrightarrow{AA_i} = d\boldsymbol{r}$ we have, according to figure 5

$$d\boldsymbol{r} = -\boldsymbol{r} + \overrightarrow{EE_i} + \boldsymbol{r}'$$
$$d\boldsymbol{r} = -\boldsymbol{r} + \boldsymbol{u}\left(\frac{r}{c} - \frac{r'}{c}\right) + \boldsymbol{r}'$$
$$d\boldsymbol{r} = -\frac{\boldsymbol{\upsilon}}{c}(r - r') + u\left(\frac{r}{c} - \frac{r'}{c}\right)$$

$$d\boldsymbol{r} = -(r - r')\left(\frac{\boldsymbol{\upsilon}}{c} - \frac{\boldsymbol{u}}{c}\right). \tag{63}$$

For the time interval dw, during which the material particle moved from point E to point E_i, it is

$$dw = \left(t - \frac{r'}{c}\right) - \left(t - \frac{r}{c}\right) = \frac{r}{c} - \frac{r'}{c},$$

therefore from equation (63) we obtain

$$\overrightarrow{AA_i} = d\boldsymbol{r} = -c\, dw\left(\frac{\boldsymbol{\upsilon}}{c} - \frac{\boldsymbol{u}}{c}\right). \tag{64}$$

In figure 5 we consider curve C which includes all the generalized photons emitted by the material particle during the time interval from $w = t - \frac{r}{c}$ to t towards a particular direction $\frac{\boldsymbol{\upsilon}}{c}$, that is, with the same velocity $\boldsymbol{\upsilon}$.

We now consider the tangent vector t [4] of the curve C at point A

$$\boldsymbol{t} = \frac{d\boldsymbol{r}}{\|d\boldsymbol{r}\|} = \frac{\frac{\boldsymbol{u}}{c} - \frac{\boldsymbol{\upsilon}}{c}}{\left\|\frac{\boldsymbol{u}}{c} - \frac{\boldsymbol{\upsilon}}{c}\right\|} = \frac{\boldsymbol{u} - \boldsymbol{\upsilon}}{\|\boldsymbol{u} - \boldsymbol{\upsilon}\|} \tag{65}$$

as follows from equation (64). For the three-dimensional arc length dS of curve C at point A we obtain from equation (64)

$$dS = \|d\boldsymbol{r}\| = dw\,\|\boldsymbol{u} - \boldsymbol{\upsilon}\|. \tag{66}$$

Now, we calculate the curvature k and the torsion τ of curve C at point A. First, we calculate the curvature vector $\boldsymbol{k}$

$$\boldsymbol{k} = \frac{d\boldsymbol{t}}{ds} = \frac{d\boldsymbol{t}}{dw\,\|\boldsymbol{u} - \boldsymbol{\upsilon}\|} = \frac{1}{\|\boldsymbol{u} - \boldsymbol{\upsilon}\|}\frac{d}{dw}\left(\frac{\boldsymbol{u} - \boldsymbol{\upsilon}}{\|\boldsymbol{u} - \boldsymbol{\upsilon}\|}\right). \tag{67}$$

Taking into account that $\frac{d\boldsymbol{\upsilon}}{dw} = 0$, $\frac{d\boldsymbol{u}}{dw} = \boldsymbol{\alpha}$ and $\|\boldsymbol{u} - \boldsymbol{\upsilon}\| = \sqrt{c^2 + \boldsymbol{u}^2 - 2(\boldsymbol{\upsilon}\cdot\boldsymbol{u})}$, we calculate the vector

$$\boldsymbol{n} = \frac{\boldsymbol{k}}{\|\boldsymbol{k}\|} = \frac{(\boldsymbol{u} - \boldsymbol{\upsilon}) \times [\boldsymbol{\alpha} \times (\boldsymbol{u} - \boldsymbol{\upsilon})]}{\|\boldsymbol{u} - \boldsymbol{\upsilon}\|\,\|(\boldsymbol{u} - \boldsymbol{\upsilon}) \times \boldsymbol{\alpha}\|}. \tag{68}$$

Combining equations (65) and (68), we calculate vector $\boldsymbol{b} = \boldsymbol{t} \times \boldsymbol{n}$ appearing in the Frenet formulas:

$$\boldsymbol{b} = \frac{(\boldsymbol{u} - \boldsymbol{\upsilon}) \times \boldsymbol{\alpha}}{\|(\boldsymbol{u} - \boldsymbol{\upsilon}) \times \boldsymbol{\alpha}\|}. \tag{69}$$

We remind that the Frenet equations

$$\begin{aligned} \frac{d\boldsymbol{t}}{ds} &= k\boldsymbol{n} \\ \frac{d\boldsymbol{n}}{ds} &= -k\boldsymbol{t}+\tau\boldsymbol{b} \\ \frac{d\boldsymbol{b}}{ds} &= -\tau\boldsymbol{n} \end{aligned} \tag{70}$$

uniquely determine the curve C. Having calculated vectors $\boldsymbol{t}, \boldsymbol{n}, \boldsymbol{b}$ we now determine the curvature k and the torsion τ of curve C from equations (70). After the necessary calculations, we obtain

$$k = \frac{\sqrt{\|\boldsymbol{u}-\boldsymbol{v}\|^2 \|\boldsymbol{\alpha}\|^2 - [\boldsymbol{\alpha}\cdot(\boldsymbol{u}-\boldsymbol{v})]^2}}{\|\boldsymbol{u}-\boldsymbol{v}\|^3} \tag{71}$$

$$\tau = \frac{\boldsymbol{\alpha}\left[(\boldsymbol{u}-\boldsymbol{v})\times\dfrac{d\boldsymbol{\alpha}}{dw}\right]}{\|\boldsymbol{\alpha}\|^2 \|\boldsymbol{u}-\boldsymbol{v}\|^2 - [(\boldsymbol{u}-\boldsymbol{v})\cdot\boldsymbol{\alpha}]^2} \|\boldsymbol{u}-\boldsymbol{v}\|^2 . \tag{72}$$

We repeat the same procedure deriving vectors $\boldsymbol{t}_p, \boldsymbol{k}_p$ and $\boldsymbol{b}_p$ at point E of the curve C_p of the material particle. For $\|\boldsymbol{u}\| \neq 0$ it is

$$\boldsymbol{t}_p = \frac{\boldsymbol{u}}{\|\boldsymbol{u}\|} \tag{73}$$

while the three-dimensional arc length is

$$dS_p = \|\boldsymbol{u}\|\, dw. \tag{74}$$

The curvature vector $\boldsymbol{k}_p$ is given by

$$\boldsymbol{k}_p = \frac{d\boldsymbol{t}_p}{dS_p} = \frac{1}{\|\boldsymbol{u}\|}\frac{d}{dw}\left(\frac{\boldsymbol{u}}{\|\boldsymbol{u}\|}\right) = \frac{\boldsymbol{\alpha}}{\|\boldsymbol{u}\|^2} - \frac{(\boldsymbol{u}\cdot\boldsymbol{\alpha})}{\|\boldsymbol{u}\|^4}\boldsymbol{u}$$

and finally,

$$\boldsymbol{k}_p = \frac{\boldsymbol{u}\times(\boldsymbol{\alpha}\times u)}{\|u\|^4}. \tag{75}$$

From equation (75) we get for vector $\boldsymbol{n}_p$

$$\boldsymbol{n}_p = \frac{\boldsymbol{k}_p}{\left\|\boldsymbol{k}_p\right\|} = \frac{\boldsymbol{u}\times(\boldsymbol{\alpha}\times u)}{\|\boldsymbol{u}\|\,\|\boldsymbol{\alpha}\times u\|}. \tag{76}$$

From equations (73) and (76) we get vector

$$\boldsymbol{b}_p = \boldsymbol{t}_p \times \boldsymbol{n}_p$$

$$\boldsymbol{b}_p = \frac{\boldsymbol{u}\times\boldsymbol{\alpha}}{\|\boldsymbol{u}\times\boldsymbol{\alpha}\|}. \tag{77}$$

From the Frenet formulas (70) for curve C_p, we get for the curvature k_p and the torsion τ_p:

$$k_p = \frac{\sqrt{\|\boldsymbol{u}\|^2 \|\boldsymbol{\alpha}\|^2 - (\boldsymbol{u}\cdot\boldsymbol{\alpha})^2}}{\|\boldsymbol{u}\|^3} \tag{78}$$

$$\tau_p = \frac{\boldsymbol{\alpha}\cdot\left(\boldsymbol{u}\times\dfrac{d\boldsymbol{\alpha}}{dw}\right)}{\|\boldsymbol{\alpha}\|^2 \|\boldsymbol{u}\|^2 - (\boldsymbol{u}\cdot\boldsymbol{\alpha})^2} \|\boldsymbol{u}\|^2 . \tag{79}$$

Comparing equations (65), (68), (69), (71) and (72) for curve C, with equations (73), (76), (77), (78) and (79) for curve C_p we arrive at the following theorem:

Theorem 1. *Trajectory representation theorem.*
For every direction $\frac{\boldsymbol{v}}{c}$ the following hold:

1. *The map $f : \boldsymbol{u} \to \boldsymbol{u}-\boldsymbol{v}$ maps the trajectory C_p of the material particle to the curve C of the generalized photons moving with velocity $\boldsymbol{v}$*

$$f : \left(\boldsymbol{t}_p, \boldsymbol{n}_p, \boldsymbol{b}_p, k_p, \tau_p\right) \to (\boldsymbol{t}, \boldsymbol{n}, \boldsymbol{b}, k, \tau) .$$

2. *The map $f^{-1} : \boldsymbol{u}-\boldsymbol{v} \to \boldsymbol{u}$ maps the curve C of the generalized photons moving with velocity $\boldsymbol{v}$ to the curve C_p of the material particle:*

$$f^{-1} : (\boldsymbol{t}, \boldsymbol{n}, \boldsymbol{b}, k, \tau) \to \left(\boldsymbol{t}_p, \boldsymbol{n}_p, \boldsymbol{b}_p, k_p, \tau_p\right).$$

According to the theorem (1), if we know the position $P(x, y, z, t)$ of the material particle at moment t and the trajectory C_p at some past time, we can determine the distribution of the generalized photons the material particle has emitted in this specific past time. We know exactly how each kinematic characteristic of the material particle maps to its surrounding spacetime.

2.5 The fundamental mathematical theorem

The interaction of the material point particle with the surrounding spacetime depends on the following four parameters:

- The moment $w = t - \frac{r}{c}$ of emission of the generalized photon by the material particle. All the physical quantities, such as the rest mass, the electric charge, the velocity $\boldsymbol{u} = \boldsymbol{u}(w)$ and the acceleration $\boldsymbol{\alpha} = \alpha(w)$ of the material particle depend upon the moment w of the emission of the generalized photon.
- The distance $r = \|\boldsymbol{r}\|$ of the arbitrary point $A(x, y, z, t)$, as depicted in figure 4, from the point of emission

$$E\left(x_p(w), y_p(w), z_p(w), w\right)$$

of the generalized photon.
- The direction in space, i.e. the functions $\delta = \delta(x, y, z, t)$ and $\omega = \omega(x, y, z, t)$.

In this paragraph we will prove the fundamental equations concerning these four parameters.

Initially we prove that the vectors ∇w, $\nabla\delta$ and $\nabla\omega$ are linearly independent. Let us suppose that

$$\lambda_1\nabla w + \lambda_2\nabla\delta + \lambda_3\nabla\omega = 0, \lambda_1, \lambda_2, \lambda_2 \in \mathbb{R}.$$

Taking into account equations (12), (51) and (52), we obtain

$$-\lambda_1\frac{1}{c\left(1-\frac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}\right)}\frac{\boldsymbol{v}}{c}+\lambda_2\left(\frac{\boldsymbol{u}\cdot\boldsymbol{\beta}}{r\left(1-\frac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}\right)}\frac{\boldsymbol{v}}{c}+\frac{1}{r}\boldsymbol{\beta}\right)+$$
$$+\lambda_3\left(\frac{\boldsymbol{u}\cdot\boldsymbol{\gamma}}{\sin\delta r\left(1-\frac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}\right)}\frac{\boldsymbol{v}}{c}+\frac{1}{r\sin\delta}\boldsymbol{\gamma}\right)=0.$$

From the linear independence of the vectors $\frac{\boldsymbol{v}}{c},\boldsymbol{\beta},\boldsymbol{\gamma}$ we see that

$$\frac{-\lambda_1}{c}+\lambda_2\frac{\boldsymbol{u}\cdot\boldsymbol{\beta}}{r}+\lambda_3\frac{\boldsymbol{u}\cdot\boldsymbol{\gamma}}{r\sin\delta}=0$$
$$\frac{\lambda_2}{r}=0$$
$$\frac{\lambda_3}{r\sin\delta}=0.$$

Finally, we have $\lambda_1=\lambda_2=\lambda_3=0$. Therefore the vectors $\nabla w,\nabla\delta,\nabla\omega$ are linearly independent.

We now focus our attention on the variation of the quantities w,δ,ω and r on the trajectory of the material particle and on the trajectory of the generalized photon. The following two theorems hold:

Theorem 2.

$$\frac{\partial w}{\partial t}+\boldsymbol{u}\cdot\nabla w=1 \tag{80a}$$
$$\frac{\partial\delta}{\partial t}+\boldsymbol{u}\cdot\nabla\delta=0 \tag{80b}$$
$$\frac{\partial\omega}{\partial t}+\boldsymbol{u}\cdot\nabla\omega=0 \tag{80c}$$
$$\frac{\partial r}{\partial t}+\boldsymbol{u}\nabla r=0. \tag{80d}$$

Theorem 3.

$$\frac{\partial w}{\partial t}+\boldsymbol{v}\cdot\nabla w=0 \tag{81a}$$
$$\frac{\partial\delta}{\partial t}+\boldsymbol{v}\cdot\nabla\delta=0 \tag{81b}$$
$$\frac{\partial\omega}{\partial t}+\boldsymbol{v}\cdot\nabla\omega=0 \tag{81c}$$
$$\frac{\partial r}{c\partial t}+\frac{\boldsymbol{v}}{c}\cdot\nabla r=1. \tag{81d}$$

From equations (11) and (12) we have

$$\frac{\partial w}{\partial t}+\boldsymbol{u}\cdot\nabla w=\frac{1}{1-\frac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}}-\frac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2\left(1-\frac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}\right)}=\frac{1-\frac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}}{1-\frac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}}=1$$
$$\frac{\partial w}{\partial t}+\boldsymbol{v}\cdot\nabla w=\frac{1}{1-\frac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}}-\frac{\|\boldsymbol{v}\|^2}{c^2\left(1-\frac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}\right)}=0.$$

From equations (41) and (51) we have

$$\frac{\partial\delta}{\partial t}+\boldsymbol{u}\nabla\delta=-\frac{\boldsymbol{u}\cdot\boldsymbol{\beta}}{r\left(1-\frac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}\right)}+\boldsymbol{u}\left(\frac{(\boldsymbol{u}\cdot\boldsymbol{\beta})}{r\left(1-\frac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}\right)}\frac{\boldsymbol{v}}{c^2}+\frac{1}{r}\boldsymbol{\beta}\right)$$
$$=-\frac{\boldsymbol{u}\cdot\boldsymbol{\beta}}{r\left(1-\frac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}\right)}+\frac{(\boldsymbol{u}\cdot\boldsymbol{\beta})\left(\boldsymbol{u}\frac{\boldsymbol{v}}{c^2}\right)}{r\left(1-\frac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}\right)}+\frac{\boldsymbol{u}\cdot\boldsymbol{\beta}}{r}=$$
$$=\frac{\boldsymbol{u}\cdot\boldsymbol{\beta}}{r\left(1-\frac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}\right)}\left[-1+\frac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}+1-\frac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}\right]=0.$$

$$\frac{\partial\delta}{\partial t}+\boldsymbol{v}\cdot\nabla\delta=-\frac{\boldsymbol{u}\cdot\boldsymbol{\beta}}{r\left(1-\frac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}\right)}+\boldsymbol{v}\left(\frac{(\boldsymbol{u}\cdot\boldsymbol{\beta})}{r\left(1-\frac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}\right)}\frac{\boldsymbol{v}}{c^2}+\frac{1}{r}\boldsymbol{\beta}\right)$$
$$=-\frac{\boldsymbol{u}\cdot\boldsymbol{\beta}}{r\left(1-\frac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}\right)}+\frac{(\boldsymbol{u}\cdot\boldsymbol{\beta})}{r\left(1-\frac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}\right)}\frac{\|\boldsymbol{v}\|^2}{c^2}+\frac{\boldsymbol{v}\cdot\boldsymbol{\beta}}{r}=0$$

since $\|\boldsymbol{v}\|^2=c^2$ and $\boldsymbol{v}\cdot\boldsymbol{\beta}=0$.

Similarly, starting from equations (42) and (52) we arrive at equations (80)(c) and (81)(c).

From equations (9) and (10) we get

$$\frac{\partial r}{c\partial t}+\frac{\boldsymbol{u}}{c}\nabla r=-\frac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2\left(1-\frac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}\right)}+\frac{\boldsymbol{u}}{c}\left(\frac{1}{1-\frac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}}\frac{\boldsymbol{v}}{c}\right)=0$$
$$\frac{\partial r}{c\partial t}+\frac{\boldsymbol{v}}{c}\nabla r=-\frac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2\left(1-\frac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}\right)}+\frac{\boldsymbol{v}}{c}\left(\frac{1}{1-\frac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}}\frac{\boldsymbol{v}}{c}\right)=$$
$$=-\frac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2\left(1-\frac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}\right)}+\frac{1}{1-\frac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}}=$$
$$\frac{1}{1-\frac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}}\left(1-\frac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}\right)=1.$$

With the aid of the above theorems we can prove the following fundamental theorem:

Theorem 4. *The Fundamental Mathematical Theorem. For every function $f=f(w,\delta,\omega,r)$ the following hold:*

A)

$$\frac{\partial f}{\partial t}+\boldsymbol{u}\cdot\nabla f=\frac{\partial f}{\partial w} \tag{82}$$
$$\frac{\partial}{\partial t}\left(f\frac{\boldsymbol{v}}{c}\right)+\left(\mathrm{grad}\left(\mathrm{f}\frac{\boldsymbol{v}}{\mathrm{c}}\right)\right)\boldsymbol{u}=\frac{\upsilon}{c}\frac{\partial f}{\partial w} \tag{83}$$
$$\frac{\partial}{\partial t}(f\boldsymbol{\beta})+(\mathrm{grad}\,(\mathrm{f}\boldsymbol{\beta}))\,\boldsymbol{u}=\boldsymbol{\beta}\frac{\partial f}{\partial w} \tag{84}$$

$$\frac{\partial}{\partial t}(f\boldsymbol{\gamma})+(\operatorname{grad}(\mathrm{f}\boldsymbol{\gamma}))\boldsymbol{u}=\boldsymbol{\gamma}\frac{\partial f}{\partial w}. \tag{85}$$

B)

$$\frac{\partial f}{\partial t}+\boldsymbol{\upsilon}\cdot\nabla f=c\frac{\partial f}{\partial r} \tag{86}$$

$$\frac{\partial}{\partial t}\left(f\frac{\boldsymbol{\upsilon}}{c}\right)+\left(\operatorname{grad}\left(\mathrm{f}\frac{\boldsymbol{\upsilon}}{\mathrm{c}}\right)\right)\boldsymbol{\upsilon}=\boldsymbol{\upsilon}\frac{\partial f}{\partial r} \tag{87}$$

$$\frac{\partial}{\partial t}(f\boldsymbol{\beta})+(\operatorname{grad}(\mathrm{f}\boldsymbol{\beta}))\boldsymbol{\upsilon}=\boldsymbol{\beta}\frac{\partial f}{\partial r} \tag{88}$$

$$\frac{\partial}{\partial t}(f\boldsymbol{\gamma})+(\operatorname{grad}(\mathrm{f}\boldsymbol{\gamma}))\boldsymbol{\upsilon}=\boldsymbol{\gamma}\frac{\partial f}{\partial r}. \tag{89}$$

We prove equations (82), (83) and (86). The rest of the equations of the fundamental mathematical theorem are proven similarly. For the proof of equation (82) we have

$$\begin{aligned}&\frac{\partial f}{\partial t}+\boldsymbol{u}\cdot\nabla f=\frac{\partial f}{\partial w}\frac{\partial w}{\partial t}+\frac{\partial f}{\partial\delta}\frac{\partial\delta}{\partial t}+\frac{\partial f}{\partial\omega}\frac{\partial\omega}{\partial t}+\frac{\partial f}{\partial r}\frac{\partial r}{\partial t}\\&+\boldsymbol{u}\cdot\left(\frac{\partial f}{\partial w}\nabla w+\frac{\partial f}{\partial\delta}\nabla\delta+\frac{\partial f}{\partial\omega}\nabla\omega+\frac{\partial f}{\partial r}\nabla r\right)\\&=\frac{\partial f}{\partial w}\left(\frac{\partial w}{\partial t}+\boldsymbol{u}\nabla w\right)+\frac{\partial f}{\partial\delta}\left(\frac{\partial\delta}{\partial t}+\boldsymbol{u}\cdot\nabla\delta\right)\\&+\frac{\partial f}{\partial\omega}\left(\frac{\partial\omega}{\partial t}+\boldsymbol{u}\cdot\nabla\omega\right)+\frac{\partial f}{\partial r}\left(\frac{\partial r}{\partial t}+\boldsymbol{u}\cdot\nabla r\right)\end{aligned}$$

and taking into account equations (80) we obtain

$$\frac{\partial f}{\partial t}+\boldsymbol{u}\cdot\nabla f=\frac{\partial f}{\partial w},$$

which is equation (82).

In order to prove equation (83) we use the identity

$$\operatorname{grad}(\mathrm{f}\boldsymbol{\alpha})=\nabla\mathrm{f}\otimes\boldsymbol{\alpha}+\mathrm{f}\ \operatorname{grad}\boldsymbol{\alpha} \tag{90}$$

which holds for every vector $\boldsymbol{\alpha}$ and scalar function f. We can now prove equation (83) as:

$$\begin{aligned}&\frac{\partial}{\partial t}\left(f\frac{\boldsymbol{\upsilon}}{c}\right)+\left(\operatorname{grad}\left(\mathrm{f}\frac{\boldsymbol{\upsilon}}{\mathrm{c}}\right)\right)\boldsymbol{u}=\\&\frac{\partial f}{\partial t}\frac{\boldsymbol{\upsilon}}{c}+f\frac{\partial}{\partial t}\left(\frac{\boldsymbol{\upsilon}}{c}\right)+\left(f\operatorname{grad}\frac{\boldsymbol{\upsilon}}{\mathrm{c}}+\nabla\mathrm{f}\otimes\frac{\boldsymbol{\upsilon}}{\mathrm{c}}\right)\boldsymbol{u}.\end{aligned}$$

Using identity (45) $(\boldsymbol{\alpha}\otimes\boldsymbol{b})\,\boldsymbol{c}=(\boldsymbol{\alpha}\cdot\boldsymbol{c})\,\boldsymbol{b}$ we obtain

$$\begin{aligned}&\frac{\partial f}{\partial t}\frac{\boldsymbol{\upsilon}}{c}+f\frac{\partial}{\partial t}\left(\frac{\boldsymbol{\upsilon}}{c}\right)+\left(f\operatorname{grad}\frac{\boldsymbol{\upsilon}}{\mathrm{c}}\right)\boldsymbol{u}+(\boldsymbol{u}\cdot\nabla f)\frac{\boldsymbol{\upsilon}}{c}=\\&\left(\frac{\partial f}{\partial t}+\boldsymbol{u}\cdot\nabla f\right)\frac{\boldsymbol{\upsilon}}{c}+f\left(\frac{\partial}{\partial t}\left(\frac{\boldsymbol{\upsilon}}{c}\right)+\left(\operatorname{grad}\frac{\boldsymbol{\upsilon}}{\mathrm{c}}\right)\boldsymbol{u}\right)=\\&\frac{\partial f}{\partial w}\frac{\boldsymbol{\upsilon}}{c}\end{aligned}$$

since $\frac{\partial f}{\partial t}+\boldsymbol{u}\cdot\nabla f=\frac{\partial f}{\partial w}$, according to equation (82) and furthermore

$$\begin{aligned}&\frac{\partial}{\partial t}\left(\frac{\boldsymbol{\upsilon}}{c}\right)+\left(\operatorname{grad}\frac{\boldsymbol{\upsilon}}{\mathrm{c}}\right)\boldsymbol{u}=\\&\frac{\partial\delta}{\partial t}\boldsymbol{\beta}+\sin\delta\frac{\partial\omega}{\partial t}\boldsymbol{\gamma}+(\nabla\delta\otimes\boldsymbol{\beta}+\sin\delta\nabla\omega\otimes\boldsymbol{\gamma})\,\boldsymbol{u}\end{aligned}$$

according to equations (38)(b),(d). Hence we obtain

$$\begin{aligned}&\frac{\partial}{\partial t}\left(\frac{\boldsymbol{\upsilon}}{c}\right)+\left(\operatorname{grad}\frac{\boldsymbol{\upsilon}}{\mathrm{c}}\right)\boldsymbol{u}=\\&\frac{\partial\delta}{\partial t}\boldsymbol{\beta}+\sin\delta\frac{\partial\omega}{\partial t}\boldsymbol{\gamma}+(\boldsymbol{u}\cdot\nabla\delta)\beta+\sin\delta\,(\boldsymbol{u}\cdot\nabla\omega)\,\boldsymbol{\gamma}=\\&\left(\frac{\partial\delta}{\partial t}+\boldsymbol{u}\cdot\nabla\delta\right)\boldsymbol{\beta}+\sin\delta\left(\frac{\partial\omega}{\partial t}+\boldsymbol{u}\cdot\nabla\omega\right)\boldsymbol{\gamma}=0\end{aligned}$$

according to equations (80)(b),(c).

The proof of equation (86) goes as follows:

$$\begin{aligned}&\frac{\partial f}{\partial t}+\boldsymbol{\upsilon}\cdot\nabla f=\frac{\partial f}{\partial w}\frac{\partial w}{\partial t}+\frac{\partial f}{\partial\delta}\frac{\partial\delta}{\partial t}+\frac{\partial f}{\partial\omega}\frac{\partial\omega}{\partial t}+\frac{\partial f}{\partial r}\frac{\partial r}{\partial t}\\&+\boldsymbol{\upsilon}\cdot\left(\frac{\partial f}{\partial w}\nabla w+\frac{\partial f}{\partial\delta}\nabla\delta+\frac{\partial f}{\partial\omega}\nabla\omega+\frac{\partial f}{\partial r}\nabla r\right)\\&=\frac{\partial f}{\partial w}\left(\frac{\partial w}{\partial t}+\boldsymbol{\upsilon}\cdot\nabla w\right)+\frac{\partial f}{\partial\delta}\left(\frac{\partial\delta}{\partial t}+\boldsymbol{\upsilon}\cdot\nabla\delta\right)\\&+\frac{\partial f}{\partial\omega}\left(\frac{\partial\omega}{\partial t}+\boldsymbol{\upsilon}\cdot\nabla\omega\right)+\frac{\partial f}{\partial r}\left(\frac{\partial r}{\partial t}+\boldsymbol{\upsilon}\cdot\nabla r\right).\end{aligned}$$

Taking into consideration equations (81) we get

$$\frac{\partial f}{\partial t}+\boldsymbol{\upsilon}\cdot\nabla f=c\frac{\partial f}{\partial r},$$

which is equation (86).

An immediate consequence of the theorem (4) is the following lemma:

For every vector function $\boldsymbol{F}=\boldsymbol{F}(w,\delta,\omega,r)$ the following relations hold:

$$\frac{\partial\boldsymbol{F}}{\partial t}+(\operatorname{grad}\boldsymbol{F})\cdot\boldsymbol{u}=\frac{\partial\boldsymbol{F}}{\partial w} \tag{91}$$

$$\frac{\partial\boldsymbol{F}}{\partial t}+(\operatorname{grad}\boldsymbol{F})\,\boldsymbol{\upsilon}=c\frac{\partial\boldsymbol{F}}{\partial r}. \tag{92}$$

The proof is done by writing the vector function $\boldsymbol{F}$ in the form

$$\boldsymbol{F}=F_1(w,\delta,\omega,r)\frac{\boldsymbol{\upsilon}}{c}+F_2(w,\delta,\omega,r)\boldsymbol{\beta}+F_3(w,\delta,\omega,r)\boldsymbol{\gamma}$$

and applying the theorem.

The fundamental mathematical theorem determines the variation of any scalar, vectorial and tensorial physical quantity, both as defined on the material particle, as well as on the surrounding spacetime. Of special interest are the applications of this theorem for the variations of the rest mass, the

electric charge, the energy, the linear momentum, the angular momentum, and any other conserved physical quantity, for the system "material particle-generalized photon". The fundamental theorem allows us to correlate the variations that take place on the material particle with the corresponding variations that take place in the surrounding spacetime.

2.6 The properties of the vector basis $\frac{\upsilon}{c}, \beta, \gamma$

The properties of the right-handed orthonormal vector basis $\{\frac{\upsilon}{c}, \beta, \gamma\}$ are given by equations (38), (39) and (40). In these equations we already know their second parts from the study conducted in the preceding paragraphs. Thus, we can express them in a simpler form.

The first of equations (38), (39) and (40) can be written as:

$$\nabla \cdot \left(\frac{\upsilon}{c}\right) = \frac{2}{r} \tag{93}$$

$$\nabla \cdot \beta = -\frac{u \cdot \beta}{cr\left(1 - \dfrac{\upsilon \cdot u}{c^2}\right)} + \frac{\cos\delta}{r \sin\delta} \tag{94}$$

$$\nabla \cdot \gamma = -\frac{u \cdot \gamma}{cr\left(1 - \dfrac{\upsilon \cdot u}{c^2}\right)}. \tag{95}$$

Equation (93) results directly from equation (22). But we can also prove it in a different way, starting from the first of equations (38)

$$\nabla \cdot \left(\frac{\upsilon}{c}\right) = \beta \cdot \nabla\delta + \sin\delta\gamma \cdot \nabla\omega.$$

With the help of equations (51) and (52) we obtain

$$\nabla \cdot \left(\frac{\upsilon}{c}\right) = \frac{1}{r} + \frac{1}{r} = \frac{2}{r}$$

taking into account that the set of the vectors $\{\frac{\upsilon}{c}, \beta, \gamma\}$ form a right-handed, orthonormal vector basis.

From the first of equations (39) we obtain

$$\nabla \cdot \beta = -\frac{\upsilon}{c}\nabla\delta + \cos\delta\gamma \cdot \nabla\omega.$$

Through equations (51) and (52) we get

$$\nabla \cdot \beta = -\frac{u \cdot \beta}{cr\left(1 - \dfrac{\upsilon \cdot u}{c^2}\right)} + \frac{\cos\delta}{r\sin\delta}.$$

From the first of equations (40) we have that

$$\nabla \cdot \gamma = -\sin\delta\frac{\upsilon}{c}\nabla\omega - \cos\delta\beta \cdot \nabla\omega.$$

Using equation (52) we see that

$$\nabla \cdot \gamma = -\frac{u \cdot \gamma}{cr\left(1 - \dfrac{\upsilon \cdot u}{c^2}\right)}.$$

Accordingly we can write in a simpler form the rest of the equations (38), (39) and (40), whenever it is demanded by the mathematical calculations performed.

2.7 List of auxiliary equations

We prove the following auxiliary equations:

$$\frac{\partial(\boldsymbol{\upsilon} \cdot \boldsymbol{u})}{\partial t} = \frac{\boldsymbol{\upsilon} \cdot \boldsymbol{\alpha}}{1 - \dfrac{\boldsymbol{\upsilon} \cdot \boldsymbol{u}}{c^2}} + \frac{(\boldsymbol{\upsilon} \cdot \boldsymbol{u})^2 - c^2u^2}{c^3 r\left(1 - \dfrac{\boldsymbol{\upsilon} \cdot \boldsymbol{u}}{c^2}\right)} \tag{96}$$

$$\nabla(\boldsymbol{\upsilon} \cdot \boldsymbol{u}) = -\frac{\boldsymbol{\upsilon} \cdot \boldsymbol{\alpha}}{c^2\left(1 - \dfrac{\boldsymbol{\upsilon} \cdot \boldsymbol{u}}{c^2}\right)}\boldsymbol{\upsilon} + \frac{c}{r}\boldsymbol{u} + \frac{u^2 - (\boldsymbol{\upsilon} \cdot \boldsymbol{u})}{cr\left(1 - \dfrac{\boldsymbol{\upsilon} \cdot \boldsymbol{u}}{c^2}\right)}\frac{\upsilon}{c} \tag{97}$$

$$\frac{\partial(\boldsymbol{\upsilon} \cdot \boldsymbol{\alpha})}{\partial t} = \frac{\boldsymbol{\upsilon} \cdot \boldsymbol{b}}{1 - \dfrac{\boldsymbol{\upsilon} \cdot \boldsymbol{u}}{c^2}} + \frac{(\boldsymbol{\upsilon} \cdot \boldsymbol{u})(\boldsymbol{\upsilon} \cdot \boldsymbol{\alpha}) - c^2(\boldsymbol{\upsilon} \cdot \boldsymbol{\alpha})}{cr\left(1 - \dfrac{\boldsymbol{\upsilon} \cdot \boldsymbol{u}}{c^2}\right)} \tag{98}$$

$$\nabla(\boldsymbol{\upsilon} \cdot \boldsymbol{\alpha}) = -\frac{\boldsymbol{\upsilon} \cdot \boldsymbol{b}}{c^2\left(1 - \dfrac{\boldsymbol{\upsilon} \cdot \boldsymbol{u}}{c^2}\right)}\boldsymbol{\upsilon} + \frac{c}{r}\boldsymbol{\alpha} + \frac{\boldsymbol{u} \cdot \boldsymbol{\alpha} - \upsilon \cdot \boldsymbol{\alpha}}{cr\left(1 - \dfrac{\boldsymbol{\upsilon} \cdot \boldsymbol{u}}{c^2}\right)}\boldsymbol{\upsilon} \tag{99}$$

where $\boldsymbol{\alpha} = \boldsymbol{\alpha}(w) = \frac{d\boldsymbol{u}(w)}{dw}$ and $\boldsymbol{b} = \boldsymbol{b}(w) = \frac{d\boldsymbol{\alpha}(w)}{dw}$ and $u^2 = \|\boldsymbol{u}\|^2$.

Indeed, it holds that

$$\frac{\partial(\boldsymbol{\upsilon} \cdot \boldsymbol{u})}{\partial t} = \boldsymbol{u}\frac{\partial \boldsymbol{\upsilon}}{\partial t} + \boldsymbol{\upsilon}\frac{\partial \boldsymbol{u}}{\partial t}$$

$$\frac{\partial(\boldsymbol{\upsilon} \cdot \boldsymbol{u})}{\partial t} = \boldsymbol{u}\frac{\partial \boldsymbol{\upsilon}}{\partial t} + \boldsymbol{\upsilon}\frac{\partial \boldsymbol{u}}{\partial w} \cdot \frac{\partial w}{\partial t}.$$

Through equations (25) and (11) we obtain

$$\frac{\partial(\boldsymbol{\upsilon} \cdot \boldsymbol{u})}{\partial t} = \boldsymbol{u}\frac{\partial \boldsymbol{\upsilon}}{\partial t} + \frac{\boldsymbol{\upsilon} \cdot \boldsymbol{\alpha}}{1 - \dfrac{\boldsymbol{\upsilon} \cdot \boldsymbol{u}}{c^2}}.$$

With the help of equation (17) we get

$$\frac{\partial(\boldsymbol{\upsilon} \cdot \boldsymbol{u})}{\partial t} = \frac{c}{r\left(1 - \dfrac{\boldsymbol{\upsilon} \cdot \boldsymbol{u}}{c^2}\right)}\left[\frac{(\boldsymbol{\upsilon} \cdot \boldsymbol{u})^2}{c^2} - u^2\right] + \frac{\boldsymbol{\upsilon} \cdot \boldsymbol{\alpha}}{1 - \dfrac{\boldsymbol{\upsilon} \cdot \boldsymbol{u}}{c^2}}$$

and performing the necessary algebraic transformations we obtain equation (96).

In order to prove equation (97) we start from the identity

$$\nabla(\boldsymbol{\upsilon} \cdot \boldsymbol{u}) = \left(\mathrm{grad}^{\mathrm{T}}\boldsymbol{\upsilon}\right)\boldsymbol{u} + \left(\mathrm{grad}^{\mathrm{T}}\boldsymbol{u}\right)\boldsymbol{\upsilon}$$

where $\mathrm{grad}^{\mathrm{T}}\boldsymbol{\upsilon}$ and $\mathrm{grad}^{\mathrm{T}}\boldsymbol{u}$ are the transpose matrices of grad $\boldsymbol{\upsilon}$ and grad $\boldsymbol{u}$.

From equations (20) and (28) we obtain

$$\nabla(\boldsymbol{\upsilon} \cdot \boldsymbol{u}) = \left[\frac{c}{r}I + \frac{1}{r\left(1 - \dfrac{\boldsymbol{\upsilon} \cdot \boldsymbol{u}}{c^2}\right)}\frac{\boldsymbol{\upsilon}}{c} \otimes (\boldsymbol{u} - \boldsymbol{\upsilon})\right]^T \boldsymbol{u} -$$

$$- \frac{1}{c\left(1 - \dfrac{\boldsymbol{\upsilon} \cdot \boldsymbol{u}}{c^2}\right)}\left(\frac{\boldsymbol{\upsilon}}{c} \otimes \boldsymbol{\alpha}\right)^T \boldsymbol{\upsilon}$$

$$\nabla(\boldsymbol{v}\cdot\boldsymbol{u})=\left[\frac{c}{r}I+\frac{1}{r\left(1-\dfrac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}\right)}(\boldsymbol{u}-\boldsymbol{v})\otimes\frac{\boldsymbol{v}}{c}\right]^T\boldsymbol{u}-$$
$$-\frac{1}{c\left(1-\dfrac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}\right)}\left(\boldsymbol{\alpha}\otimes\frac{\boldsymbol{v}}{c}\right)\boldsymbol{v}.$$

Using identity (45) we get

$$\nabla(\boldsymbol{v}\cdot\boldsymbol{u})=\frac{c}{r}\boldsymbol{u}+\frac{\boldsymbol{u}\cdot(\boldsymbol{u}-\boldsymbol{v})}{r\left(1-\dfrac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}\right)}\cdot\frac{\boldsymbol{v}}{c}-\frac{\boldsymbol{v}\cdot\boldsymbol{\alpha}}{c\left(1-\dfrac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}\right)}\frac{\boldsymbol{v}}{c}$$

which is equation (97). We can similarly prove equations (98) and (99). In order to prove the last equation we use equation (32), in exactly the same manner we used equation (28). In the same way, we can prove corresponding equations for all of the inner products such as $\boldsymbol{v}\cdot\boldsymbol{b}$, $\boldsymbol{u}\cdot\boldsymbol{\alpha}$ etc., that appear in the equations of the theory of selfvariations.

3 The study of the selfvariations for a material point particle moving with constant speed

3.1 Introduction

In this paragraph we present the study of the selfvariations for a material point particle moving with constant speed. This study was regarded as necessary for two reasons. The first is that constant-speed motion is the simplest possible and, therefore, we are studying the consequences of the selfvariations in their simplest version. The second reason is that arbitrary motion can be considered as a multitude of successive constant-speed motions.

By studying the constant-speed motion of a material particle we can derive the Lorentz-Einstein transformations for the physical quantities w, δ, ω, r that appear in the equations of the theory of selfvariations. Of special interest is the transformation of the volume of the generalized photon, which differs from the volume transformation of material particles as we know it within the framework of Special Relativity. After having studied both the arbitrary motion, as well as the constant-speed motion of the material particle, we have the knowledge necessary for advancing our study in the forthcoming paragraphs.

3.2 The case of a material point particle moving with constant speed

We consider a material point particle with rest mass m_0 and electric charge q, which moves with velocity $\boldsymbol{u}=\begin{bmatrix}u\\0\\0\end{bmatrix}$ in the inertial frame of reference $S(0,x,y,z,t)$, as depicted in figure 6.

At moment t when the material particle is at point $P(ut,0,0,t)$, the rest mass m_0 and the electric charge q of the material particle act at point $A(x,y,z,t)$ through the generalized photon that was emitted from point E and arrived at point A moving with velocity c. Therefore, the coordinates of point E are

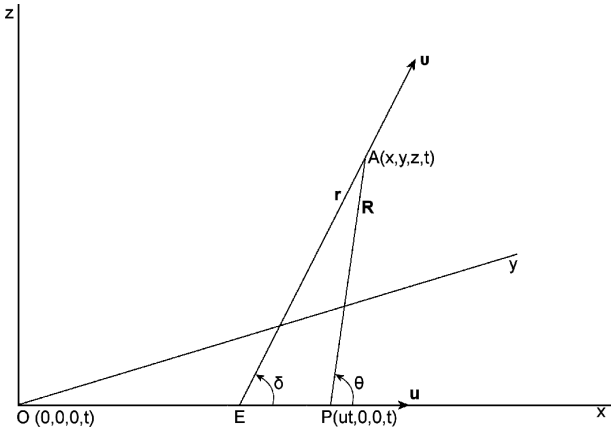

Fig. 6: Material point particle moving with constant speed along the x axis of the inertial reference frame $S(0,x,y,z,t)$. As the material particle moves from point E to point P, during the time interval $\Delta t=\frac{r}{c}$, a generalized photon moves from point E to point A.

$$E\left(ut-\frac{u}{c}r,0,0,t-\frac{r}{c}\right)\tag{100}$$

where $r=\|\boldsymbol{r}\|=\left\|\overrightarrow{EA}\right\|$. Due to the selfvariations, the rest mass m_0 and the electric charge q of the material particle act at point $A(x,y,z,t)$ with the value they had at time

$$w=t-\frac{r}{c}\tag{101}$$

at point $E\left(ut-\frac{u}{c}r,0,0,t-\frac{r}{c}\right)$, and not with the value they have at point $P(ut,0,0,t)$ at time t. For the vector $\boldsymbol{r}$ we have

$$\boldsymbol{r}=\overrightarrow{EA}=\begin{bmatrix}x-ut+\frac{u}{c}r\\y\\z\end{bmatrix}.\tag{102}$$

The magnitude of $\|\boldsymbol{r}\|=r$ can be derived from equations (102) as

$$\|\boldsymbol{r}\|=r=\gamma^2\frac{u}{c}(x-ut)+\gamma\sqrt{\gamma^2(x-ut)^2+y^2+z^2}\tag{103}$$

where $\gamma=\frac{1}{\sqrt{1-\frac{u^2}{c^2}}}$.

Combining equations (102) and (103) we obtain

$$\boldsymbol{r}=\begin{bmatrix}\gamma^2(x-ut)+\frac{u}{c}\gamma\sqrt{\gamma^2(x-ut)^2+y^2+z^2}\\y\\z\end{bmatrix}.\tag{104}$$

The velocity $\boldsymbol{v}$ of the selfvariations has magnitude $\|\boldsymbol{v}\|=c$, and is parallel to the vector $\boldsymbol{r}$, thus we have

$$\boldsymbol{v}=\frac{c}{r}\boldsymbol{r}=\frac{c}{r}\begin{bmatrix}\gamma^2(x-ut)+\frac{u}{c}\gamma\sqrt{\gamma^2(x-ut)^2+y^2+z^2}\\y\\z\end{bmatrix}.\tag{105}$$

The position vector $\boldsymbol{R}$ of point $A(x, y, z, t)$ with respect to point $P(ut, 0, 0, t)$, where the material particle is located, is

$$\boldsymbol{R} = \overrightarrow{PA} = \begin{bmatrix} x - ut \\ y \\ z \end{bmatrix}. \tag{106}$$

From equation (106) we obtain

$$\|\boldsymbol{R}\| = R = \sqrt{(x - ut)^2 + y^2 + z^2}. \tag{107}$$

From figure 6 we see that

$$\boldsymbol{r} = \overrightarrow{EA} + \boldsymbol{R}$$
$$\boldsymbol{r} = \frac{r}{c}\boldsymbol{u} + \boldsymbol{R}.$$

Finally, we obtain

$$\boldsymbol{v} = \boldsymbol{u} + \frac{c}{r}\boldsymbol{R} \tag{108}$$

$$\boldsymbol{R} = r\left(\frac{\boldsymbol{v}}{c} - \frac{\boldsymbol{u}}{c}\right). \tag{109}$$

Combining equations (100) and (101) we have for the coordinates of point E

$$E(uw, 0, 0, w). \tag{110}$$

The relations between the scalar, vectorial and tensorial quantities of this paragraph can be derived by the corresponding relations proven in the second paragraph, considering that the acceleration of the material body vanishes, that is $\boldsymbol{\alpha} = \boldsymbol{\alpha}(w) = 0$, and that the velocity of the material particle is $\boldsymbol{u} = \boldsymbol{u}(w) = \begin{bmatrix} u(w) \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} u \\ 0 \\ 0 \end{bmatrix}$.

3.3 The case of a material point particle at rest

We consider an inertial reference frame $S'(0', x', y', z', t')$ moving with velocity $\boldsymbol{u} = \begin{bmatrix} u \\ 0 \\ 0 \end{bmatrix}$ with respect to the inertial reference frame $S(0, x, y, z, t)$ of the previous paragraph. We also suppose that for $t = t' = 0$ the origins of the axes of coordinates 0 and 0′ of these two frames coincide. In the way we have chosen these two inertial frames, the material particle is at rest in frame S' or, equivalently, frame S' accompanies the material particle during its motion. Figure 7 is the one corresponding to figure 6 for reference frame S'.

At moment t', when the material particle is located at point $P(0, 0, 0, t')$, the mass m_o and the electric charge q of the material particle act at point $A(x', y', z', t')$ through the generalized photon that was emitted from point $E\left(0, 0, 0, t' - \frac{r'}{c}\right)$

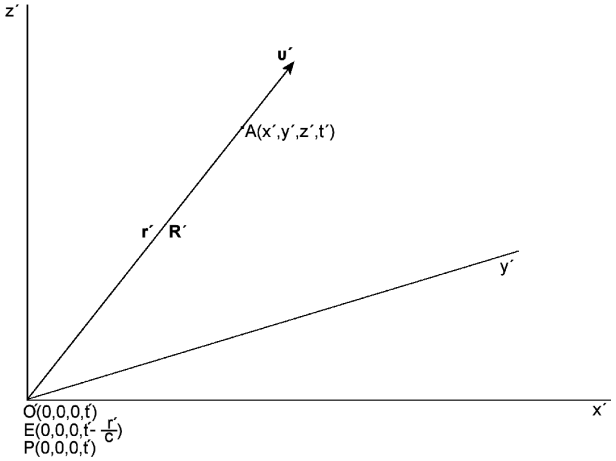

Fig. 7: A material point particle remains at rest at the origin $O'(0, 0, 0, 0, t')$ of the inertial reference frame $S(0', x', y', z', t')$. A generalized photon moves from point $E\left(0, 0, 0, 0, t' - \frac{r'}{c}\right)$ and arrives at point $A(x', y', z', t')$, during the time interval $\Delta t' = \frac{r'}{c}$.

and arrived at point $A(x', y', z', t')$ moving with velocity c. Therefore, the coordinates of point E are

$$E\left(0, 0, 0, t' - \frac{r'}{c}\right) \tag{111}$$

where $r' = \|\boldsymbol{r}'\| = \left\|\overrightarrow{EA}\right\|$. Due to the selfvariations, the rest mass m_o and the electric charge q of the material particle act at point $A(x', y', z', t')$ with the value they had at time

$$w' = t' - \frac{r'}{c} \tag{112}$$

and not with the value they have at $P(0, 0, 0, t')$.

For the vector $\boldsymbol{r}'$ it holds that

$$\boldsymbol{r}' = \overrightarrow{EA} = \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} \tag{113}$$

while its magnitude $\|\boldsymbol{r}'\| = r'$ is given by

$$\left\|\boldsymbol{r}'\right\| = r' = \sqrt{x'^2 + y'^2 + z'^2}. \tag{114}$$

The velocity of the selfvariations $\boldsymbol{v}'$ has magnitude $\|\boldsymbol{v}'\| = c$, and is parallel to the vector $\boldsymbol{r}'$, therefore it is

$$\boldsymbol{v}' = \frac{c}{r'}\boldsymbol{r}' = \frac{c}{r'}\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix}. \tag{115}$$

The position vector $\boldsymbol{R}'$ of point $A(x', y', z', t')$ with respect to $P(0, 0, 0, t')$, where the material particle is located, is given by

$$\boldsymbol{R}' = \overrightarrow{PA} = \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \boldsymbol{r}'. \tag{116}$$

From equation (116) we get

$$\left\|\boldsymbol{R}'\right\| = R' = \left\|\boldsymbol{r}'\right\| = r' = \sqrt{x'^2 + y'^2 + z'^2}. \tag{117}$$

Combining equations (111) and (112) we obtain for the coordinates of point E

$$E\,(0, 0, 0, w')\,. \tag{118}$$

The relations between the scalar, vectorial and tensorial quantities of this paragraph can be derived from the corresponding relations we proved in the second paragraph, considering that the acceleration and the velocity of the material particle vanish, that is $\boldsymbol{\alpha} = \boldsymbol{\alpha}\,(w) = \mathbf{0}$ and $\boldsymbol{u} = \boldsymbol{u}\,(w) = \mathbf{0}$.

3.4 Lorentz-Einstein transformations of the quantities w, δ, ω, r

In this paragraph we shall study the way in which the fundamental physical quantities appearing in the equations of the theory of selfvariations transform under the action of the Lorentz-Einstein transformations [5–11].

In the way we have chosen the inertial reference frames S and S', the transformations of the coordinates in the four-dimensional spacetime are given by the set of equations

$$\begin{aligned}
&x = \gamma\,(x' + ut')\\
&y = y'\\
&z = z'\\
&t = \gamma\left(t' + \frac{u}{c^2}x'\right)\\
&x' = \gamma\,(x - ut)\\
&y' = y\\
&z' = z\\
&t' = \gamma\left(t - \frac{u}{c^2}x\right)
\end{aligned} \tag{119}$$

where $\gamma = \frac{1}{\sqrt{1-\frac{u^2}{c^2}}}$.

The coordinates of point E are given by relation (110), and are $E\,(uw, 0, 0, w)$ for inertial frame S, and by relation (118), and are $E\,(0, 0, 0, w')$ for inertial frame S'. Applying transformations (119) we obtain

$$w = \gamma w'. \tag{120}$$

Indeed, based on the fourth equation of the first column of transformations (119) for the coordinates of point E, we get

$$\begin{aligned}
&w = \gamma\,(w' + u \cdot 0)\\
&w = \gamma w'.
\end{aligned}$$

We now consider the trigonometric form of the velocity $\boldsymbol{\upsilon}$, as defined in paragraph 2.2. From equations (34) we get for reference frames S and S' respectively

$$\begin{aligned}
&\cos\delta = \frac{\upsilon_x}{c}\\
&\sin\delta\cos\omega = \frac{\upsilon_y}{c}\\
&\sin\delta\sin\omega = \frac{\upsilon_z}{c}
\end{aligned} \tag{121}$$

$$\begin{aligned}
&\cos\delta' = \frac{\upsilon'_x}{c}\\
&\sin\delta'\cos\omega' = \frac{\upsilon'_y}{c}\\
&\sin\delta'\sin\omega' = \frac{\upsilon'_z}{c}.
\end{aligned} \tag{122}$$

From the Lorentz-Einstein transformations for the velocity we have

$$\begin{aligned}
&\upsilon_x = \frac{\upsilon'_x + u}{1 + \frac{u\upsilon'_x}{c^2}} \qquad \upsilon'_x = \frac{\upsilon_x - u}{1 - \frac{u\upsilon_x}{c^2}}\\
&\upsilon_y = \frac{\upsilon'_y}{\gamma\left(1 + \frac{u\upsilon'_x}{c^2}\right)} \qquad \upsilon'_y = \frac{\upsilon_y}{\gamma\left(1 - \frac{u\upsilon_x}{c^2}\right)}\\
&\upsilon_z = \frac{\upsilon'_z}{\gamma\left(1 + \frac{u\upsilon'_x}{c^2}\right)} \qquad \upsilon'_z = \frac{\upsilon_z}{\gamma\left(1 - \frac{u\upsilon'_x}{c^2}\right)}.
\end{aligned} \tag{123}$$

From transformation (123) and from equations (121) and (122) the following transformations are derived for the functions $\delta = \delta\,(x, y, z, t)$ and $\omega = \omega\,(x, y, z, t)$:

$$\begin{aligned}
&\cos\delta' = \frac{\cos\delta - \frac{u}{c}}{1 - \frac{u}{c}\cos\delta} \qquad \cos\delta = \frac{\cos\delta' + \frac{u}{c}}{1 + \frac{u}{c}\cos\delta'}\\
&\sin\delta' = \frac{\sin\delta}{\gamma\left(1 - \frac{u}{c}\cos\delta\right)} \qquad \sin\delta = \frac{\sin\delta'}{\gamma\left(1 + \frac{u}{c}\cos\delta'\right)}\\
&\omega' = \omega \qquad \omega = \omega'.
\end{aligned} \tag{124}$$

We shall prove the first equation. The rest are proven similarly.

From the first equation of the second column of transformations (123) we obtain

$$\begin{aligned}
&\upsilon'_x = \frac{\upsilon_x - u}{1 - \frac{u\upsilon_x}{c^2}}\\
&\frac{\upsilon'_x}{c} = \frac{\frac{\upsilon_x}{c} - \frac{u}{c}}{1 - \frac{u}{c}\frac{\upsilon_x}{c}}.
\end{aligned}$$

Through equations (122) and (121) we get

$$\cos\delta' = \frac{\cos\delta - \frac{u}{c}}{1 - \frac{u}{c}\cos\delta}.$$

From equation (117) and transformations (119) we see that

$$r' = \sqrt{\gamma^2\,(x - ut)^2 + y^2 + z^2}. \tag{125}$$

Combining equations (103) and (125) we get

$$r = \gamma^2 \frac{u}{c}(x - ut) + \gamma r'$$

and since

$$\gamma(x - ut) = x'$$

from transformations (119) we obtain

$$r = \gamma \frac{u}{c} x' + \gamma r'. \tag{126}$$

From equation (115) we see that

$$\upsilon'_x = \frac{c}{r'} x'$$
$$x' = r' \frac{\upsilon'_x}{c}.$$

Substituting into equation (126) we get

$$r = \gamma \frac{u\upsilon'_x}{c^2} r' + \gamma r'$$
$$r = \gamma r' \left(1 + \frac{u\upsilon'_x}{c^2}\right).$$

From equation (122) we obtain

$$r = \gamma r' \left(1 + \frac{u}{c}\cos\delta'\right)$$

and with the help of transformations (124) we get

$$r = \gamma r' \left(1 + \frac{u}{c}\frac{\cos\delta - \frac{u}{c}}{1 - \frac{u}{c}\cos\delta}\right)$$
$$r = \gamma r' \frac{1 - \frac{u^2}{c^2}}{1 - \frac{u}{c}\cos\delta}$$
$$r = \frac{r'}{\gamma\left(1 - \frac{u}{c}\cos\delta\right)}$$

$$r' = \gamma r\left(1 - \frac{u}{c}\cos\delta\right) = \gamma r\left(1 - \frac{\boldsymbol{\upsilon}\cdot\boldsymbol{u}}{c^2}\right). \tag{127}$$

From transformations (124) we obtain

$$\sin\delta' = \frac{\sin\delta}{\gamma\left(1 - \frac{u}{c}\cos\delta\right)}$$
$$\cos\delta'\frac{d\delta'}{d\delta} = \frac{\cos\delta\left(1 - \frac{u}{c}\cos\delta\right) - \sin\delta\frac{u}{c}\sin\delta}{\gamma\left(1 - \frac{u}{c}\cos\delta\right)^2}$$
$$\cos\delta'\frac{d\delta'}{d\delta} = \frac{\cos\delta - \frac{u}{c}}{\gamma\left(1 - \frac{u}{c}\cos\delta\right)^2}$$
$$\frac{\cos\delta - \frac{u}{c}}{1 - \frac{u}{c}\cos\delta}\frac{d\delta'}{d\delta} = \frac{\cos\delta - \frac{u}{c}}{\gamma\left(1 - \frac{u}{c}\cos\delta\right)^2}$$
$$\frac{d\delta'}{d\delta} = \frac{1}{\gamma\left(1 - \frac{u}{c}\cos\delta\right)}$$

$$d\delta' = \frac{1}{\gamma\left(1 - \frac{u}{c}\cos\delta\right)} d\delta. \tag{128}$$

Repeating the same procedure we also arrive at relation

$$\frac{\partial}{\partial\delta'} = \gamma\left(1 - \frac{u}{c}\cos\delta\right)\frac{\partial}{\partial\delta} \tag{129}$$

among the operators $\frac{\partial}{\partial\delta'}$ and $\frac{\partial}{\partial\delta}$.

From equation (109) we get

$$R = r\sqrt{1 + \frac{u^2}{c^2} - 2\frac{\boldsymbol{\upsilon}\cdot\boldsymbol{u}}{c^2}}$$

$$R = r\sqrt{1 + \frac{u^2}{c^2} - 2\frac{u}{c}\cos\delta}. \tag{130}$$

From equation (130) we are able, whenever it is necessary, to derive the Lorentz-Einstein transformation of the quantity R through the use of transformations (124) and (127).

We consider now the angle θ between the vectors $\boldsymbol{R}$ and $\boldsymbol{u}$, as depicted in figure 6. From the law of sines for the triangle EAP we have that

$$\frac{\sin\vartheta}{r} = \frac{\sin\delta}{R}$$
$$\sin\vartheta = \frac{r}{R}\sin\delta.$$

Using equation (130) we obtain

$$\sin\vartheta = \frac{\sin\delta}{\sqrt{1 + \frac{u^2}{c^2} - 2\frac{u}{c}\cos\delta}}. \tag{131}$$

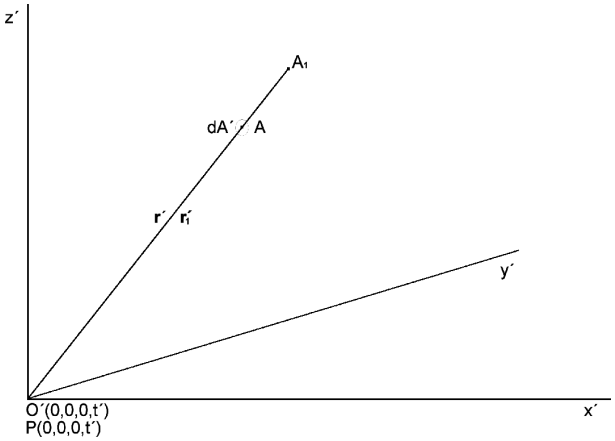

Fig. 8: The infinitesimal volume of the generalized photon in the vicinity of point A of the inertial reference frame $S\,(0', x', y', z', t')$. The material point particle is at position $P\,(0, 0, 0, t')$. The infinitesimal surface of area dA' is vertical to the vectors $r' = \vec{PA}$ and $r'_1 = \vec{P_1A}$. The points P, A and A_1 are collinear.

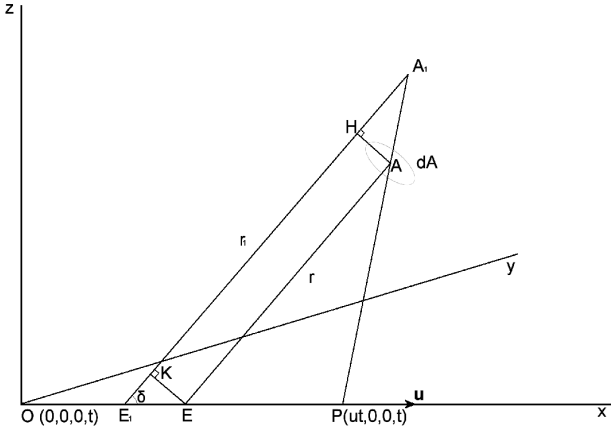

Fig. 9: Figure 8 as modulated in the inertial reference frame $S\,(0, x, y, z, t)$, in which the material particle moves with constant speed. The points P, A and A_1 remain collinear, as results from the Lorentz-Einstein transformations.

From the familiar identity $\sin^2\vartheta + \cos^2\vartheta = 1$ we have that

$$\cos\vartheta = \frac{\cos\delta - \dfrac{u}{c}}{\sqrt{1 + \dfrac{u^2}{c^2} - 2\dfrac{u}{c}\cos\delta}}. \tag{132}$$

From transformations (124) we can, after applying equations (131) and (132), derive the Lorentz-Einstein transformations for the quantities $\sin\vartheta$ and $\cos\vartheta$. Furthermore, in the inertial reference frame S' it is $\theta' = \delta'$, as can be seen from figure 7.

3.5 The Lorentz-Einstein transformation of the volume of the generalized photon

The generalized photon moves with velocity $\boldsymbol{\upsilon}$ of magnitude $\|\boldsymbol{\upsilon}\| = c$ in any inertial reference frame. This has as a consequence that the following transformation does not hold:

$$dV' = \gamma dV.$$

This transformation holds for the volume dV of a material particle that is at rest in the inertial reference frame S'. We shall prove that the volume of the generalized photon transforms according to relation

$$dV' = \frac{dV}{\gamma\left(1 - \dfrac{u}{c}\cos\delta\right)} = \frac{dV}{\gamma\left(1 - \dfrac{\boldsymbol{\upsilon}\cdot\boldsymbol{u}}{c^2}\right)} \tag{133}$$

for our chosen inertial reference frames S and S'.

In the region of point $A\,(x', y', z', t')$ of figure 7 we consider the elementary area

$$dA' = r'^2 \sin\delta' d\delta' d\omega'$$

of a sphere with center O' and radius r'. Furthermore, we consider a point A_1 close to point A on line OA, as depicted in figure 8.

The elementary volume of the generalized photon in the inertial reference frame S' is

$$dV' = dA' \left\|\overrightarrow{AA_1}\right| = r'^2 \sin\delta' d\delta' d\omega' \left\|\overrightarrow{AA_1}\right| \tag{134}$$

assuming that $A_1 \to A$.

In figure 9 we present the volume dV occupied by the generalized photon in the inertial frame of reference S.

The elementary area dA in S is

$$dA = r^2 \sin\delta d\delta d\omega$$

while the elementary volume dV is

$$dV = dA \left\|\overrightarrow{HA_1}\right| = r^2 \sin\delta d\delta d\omega \left\|\overrightarrow{HA_1}\right\| \tag{135}$$

since $A_1 \to A$.

From the Lorentz-Einstein transformations it directly follows that points P, A, A_1, which are collinear in reference frame S' are also collinear in reference frame S. The conclusions of paragraph 2.4 about the representation of the trajectory of the material particle in the surrounding spacetime, also lead to figure 9. Here, the trajectory of the material particle is on the x axis. We now use the following notation, as depicted in figure 9.

$$r = \left\|\overrightarrow{EA}\right\| \tag{136}$$

$$r_1 = \left\|\overrightarrow{E_1A_1}\right\| \tag{137}$$

according to the notation we have established. Similarly, in figure 8 we use the notation

$$r' = \left\|\overrightarrow{O'A}\right\| \tag{138}$$

$$r'_1 = \left\|\overrightarrow{O'A_1}\right\|. \tag{139}$$

From figure 9 we have that

$$\left\|\overrightarrow{E_1A_1}\right\| = \left\|\overrightarrow{E_1K}\right\| + \left\|\overrightarrow{KH}\right\| + \left\|\overrightarrow{HA_1}\right\|$$

and with equations (136) and (137) we get

$$r_1 = \left\|\overrightarrow{E_1K}\right\| + r + \left\|\overrightarrow{HA_1}\right\|$$

$$\left\|\overrightarrow{HA_1}\right\| = r_1 - r - \left\|\overrightarrow{E_1K}\right\|. \tag{140}$$

From the triangle E_1KE of figure 9 we see that

$$\cos\delta = \frac{\left\|\overrightarrow{E_1K}\right\|}{\left\|\overrightarrow{E_1E}\right\|}$$

$$\left\|\overrightarrow{E_1K}\right\| = \left\|\overrightarrow{E_1E}\right\|\cos\delta. \tag{141}$$

Similarly, we have that

$$\left\|\overrightarrow{E_1E}\right\| = u\frac{r_1 - r}{c} = u\,dw \tag{142}$$

since in the time interval $\Delta t = \frac{r_1 - r}{c}dw$ the point particle moved from point E_1 to point E. Combining equations (141) and (142) we obtain

$$\left\|\overrightarrow{E_1K}\right\| = u\,dw\cos\delta. \tag{143}$$

Combining equations (140) and (143) we also get

$$\left\|\overrightarrow{HA_1}\right\| = c\,dw\left(1 - \frac{u}{c}\cos\delta\right) \tag{144}$$

since $r_1 - r = cdw$.

Combining equations (135) and (144) we get

$$dV = r^2\sin\delta d\delta d\omega cdw\left(1 - \frac{u}{c}\cos\delta\right). \tag{145}$$

From figure 8 we have that

$$\left\|\overrightarrow{AA_1}\right\| = \left\|\overrightarrow{O'A_1}\right\| - \left\|\overrightarrow{O'A}\right\|$$

and with equations (138) and (139) we get

$$\left\|\overrightarrow{AA_1}\right\| = r'_1 - r' = cdw'. \tag{146}$$

Combining equations (135) and (146) we also get

$$dV' = r'^2\sin\delta' d\delta' d\omega' cdw'. \tag{147}$$

Combining equations (145) and (147) we get

$$\frac{dV'}{dV} = \frac{r'^2\sin\delta' d\delta' d\omega' cdw'}{r^2\sin\delta d\delta d\omega cdw\left(1 - \frac{u}{c}\cos\delta\right)}$$

and with transformations (127), (124), (128) and (121) we get

$$\frac{dV'}{dV} = \gamma^2\left(1 - \frac{u}{c}\cos\delta\right)^2\frac{1}{\gamma^2\left(1 - \frac{u}{c}\cos\delta\right)^2}\frac{1}{\gamma}\frac{1}{1 - \frac{u}{c}\cos\delta}$$

$$\frac{dV'}{dV} = \frac{1}{\gamma\left(1 - \frac{u}{c}\cos\delta\right)}$$

$$dV' = \frac{dV}{\gamma\left(1 - \frac{u}{c}\cos\delta\right)}. \tag{148}$$

This is equation (133). Given that $\boldsymbol{u} = \begin{bmatrix} u \\ 0 \\ 0 \end{bmatrix}$ we arrive at relation

$$\frac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2} = \frac{u}{c}\frac{v_x}{c} = \frac{u}{c}\cos\delta \tag{149}$$

since, according to equation (121), $\cos\delta = \frac{v_x}{c}$.

Combining equations (148) and (149) we have

$$dV' = \frac{dV}{\gamma\left(1 - \frac{u}{c}\cos\delta\right)} = \frac{dV}{\gamma\left(1 - \frac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}\right)}.$$

This is the final form of equation (133).

In the form

$$dV' = \frac{dV}{\gamma\left(1 - \frac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}\right)} \tag{150}$$

transformation (133) also holds in the case of a material particle in arbitrary motion. In figure 5 the length of the three-dimensional arc EE_i equals $\left\|\overrightarrow{EE_i}\right\|$ at first approximation, that is, for an infinitesimal displacement of the material particle from point E to point E_i. Thus, we have exactly the situation we describe in figure 9. On the other hand, for a finite, but not infinitesimal, displacement $\overrightarrow{EE_i}$ of the material particle, the curvature $k_p(w)$ and the torsion $\tau_p(w)$ of curve C_p of figure 5 enter the transformation of the volume.

4 The study of selfvariations at macroscopic scales

4.1 Introduction

In the present paragraph we study the consequences of the selfvariations at macroscopic scales. The main conclusion we derive is the existence of energy, momentum, electric charge and electric current in the surrounding spacetime of the material particle as a direct consequence of the selfvariations. We calculate the density of energy, momentum, electric charge

and electric current in the surrounding spacetime of an arbitrarily moving material point particle.

We present the four-dimensional electromagnetic potential which is compatible with the selfvariations. An important element that emerges is the splitting of the electromagnetic potential into two individual potentials, where the first one gives the electromagnetic field that accompanies the material particle in its motion, while the second one gives the electromagnetic radiation.

We prove that the selfvariations are compatible with the principles of conservation of electric charge, energy, and momentum. This is accomplished through either direct calculation, based on the continuity equation, and also through the energy-momentum tensor of the generalized photon. These different approaches help the reader comprehend the physical reality that prevails in the surrounding spacetime of material particles.

In the preceding paragraphs we studied the generalized photon as a geometric object. In this paragraph we shall see for the first time that the generalized photon is a carrier of energy, momentum, and electric charge. The density of electric charge and electric current in the surrounding spacetime of the material particle is correlated with the electromagnetic field that accompanies the material particle in its motion. The electromagnetic radiation does not contribute to the density of electric charge and electric current.

We calculate the energy-momentum tensor for the electromagnetic field and for the generalized photon. The energy-momentum tensor describes the energy content of spacetime, but only in macroscopic scales. In microscopic scales, the energy-momentum tensor, as defined by the theory of Special Relativity, cannot describe the energy content of spacetime.

4.2 The density of electric charge and electric current in the surrounding spacetime of an electrically charged point particle

In figure 6 the electric charge q acts at point $A(x, y, z, t)$ with the value it had at point E. Thus, we have $q = q(w)$. Hence, it follows that

$$\frac{\partial q}{\partial t} = \frac{\partial q}{\partial w}\frac{\partial w}{\partial t}$$

$$\nabla q = \frac{\partial q}{\partial w}\nabla w$$

and with equations (11) and (12) we have that

$$\frac{\partial q}{\partial t} = \frac{\partial q}{\partial w}\frac{1}{1-\dfrac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}} \tag{151}$$

$$\nabla q = -\frac{\partial q}{c\partial w}\frac{1}{1-\dfrac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}}\frac{\boldsymbol{v}}{c}. \tag{152}$$

According to Special Relativity and the symbols we use in figure 6, the intensity $\boldsymbol{\varepsilon}$ of the electric field at point A is

$$\boldsymbol{\varepsilon} = \frac{\gamma q}{4\pi\varepsilon_0 r'^3}\boldsymbol{R} \tag{153}$$

where $\boldsymbol{R}$ is given by equation (106), r' by equation (117), and $\gamma = \frac{1}{\sqrt{1-\frac{u^2}{c^2}}}$. From Gauss's law [12–18] we obtain for the electric charge density ρ at point A:

$$\rho = \varepsilon_0\nabla\cdot\boldsymbol{\varepsilon}$$

$$\rho = \varepsilon_0\nabla\cdot\left(\frac{\gamma q}{4\pi\varepsilon_0 r'^3}\boldsymbol{R}\right)$$

$$\rho = \frac{q\gamma}{4\pi}\nabla\cdot\left(\frac{\boldsymbol{R}}{r'^3}\right) + \frac{\gamma}{4\pi R'^3}\boldsymbol{R}\cdot\nabla q. \tag{154}$$

We can easily prove that

$$\nabla\cdot\left(\frac{\boldsymbol{R}}{r'^3}\right) = 0. \tag{155}$$

We can avoid the calculation, if we take into account that, ignoring the selfvariations, for constant electric charge q, classical Electromagnetism predicts that $\rho = 0$ at point A. This is equivalent with equation (155).

Combining equations (154) and (155) we get

$$\rho = \frac{\gamma}{4\pi r'^3}\boldsymbol{R}\cdot\nabla q.$$

Using equation (152) we get

$$\rho = -\frac{\partial q}{c\partial w}\frac{\gamma}{4\pi r'^3\left(1-\dfrac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}\right)}\frac{\boldsymbol{v}}{c}\cdot\boldsymbol{R}.$$

After applying equation (109) we have that

$$\rho = -\frac{\partial q}{c\partial w}\frac{\gamma r}{4\pi r'^3\left(1-\dfrac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}\right)}\frac{\boldsymbol{v}}{c}\left(\frac{\boldsymbol{v}}{c}-\frac{\boldsymbol{u}}{c}\right)$$

$$\rho = -\frac{\partial q}{c\partial w}\frac{\gamma r}{4\pi r'^3}\frac{1}{\left(1-\dfrac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}\right)}\left(1-\frac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}\right)$$

$$\rho = -\frac{\partial q}{c\partial w}\frac{\gamma r}{4\pi r'^3}.$$

Using transformation (127) we get

$$\rho = -\frac{\partial q}{c\partial w}\frac{1}{4\pi\gamma^2 r^2\left(1-\dfrac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}\right)^3}. \tag{156}$$

We can derive the same equation in a different way. We will develop the second method in the next paragraph for the calculation of the density of energy D due to the selfvariations of the rest mass of the material particle, where we will not be able to use Gauss's law. The reader can easily apply the

method of the next paragraph to the electric charge, and still come up with equation (156).

The generalized photon moves with velocity $\boldsymbol{\upsilon}$, therefore the current density $\boldsymbol{j}$ is given by equation

$$\boldsymbol{j} = \rho\boldsymbol{\upsilon} \tag{157}$$

where the charge density ρ is given by equation (156). Equation (157) can also be easily inferred from Ampere's law

$$\nabla \times \boldsymbol{B} = \mu_0 \boldsymbol{j} + \frac{\partial \boldsymbol{\varepsilon}}{c^2 \partial t}. \tag{158}$$

The intensity of the magnetic field $\boldsymbol{B}$ at point A of figure 6 is given initially by the Biot-Savart law:

$$\boldsymbol{B} = \frac{\boldsymbol{u}}{c^2} \times \boldsymbol{\varepsilon}. \tag{159}$$

Combining equations (153) and (109) we get

$$\boldsymbol{\varepsilon} = \frac{\gamma q}{4\pi\varepsilon_0 r'^3} r \left(\frac{\boldsymbol{\upsilon}}{c} - \frac{u}{c} \right)$$

and from equation (127) we have

$$\boldsymbol{\varepsilon} = \frac{q}{4\pi\varepsilon_0 \gamma^2 r^2 \left(1 - \dfrac{\boldsymbol{\upsilon} \cdot \boldsymbol{u}}{c^2} \right)^3} \left(\frac{\boldsymbol{\upsilon}}{c} - \frac{u}{c} \right). \tag{160}$$

From equation (160) we get

$$\frac{\boldsymbol{u}}{c^2} \times \boldsymbol{\varepsilon} = \frac{\boldsymbol{\upsilon}}{c} \times \boldsymbol{\varepsilon}$$

and from equation (159) we get

$$\boldsymbol{B} = \frac{\boldsymbol{\upsilon}}{c^2} \times \boldsymbol{\varepsilon}. \tag{161}$$

In equation (161) the velocity $\boldsymbol{\upsilon}$ of the generalized photon refers to point A of figure 6. This has as a consequence that all physical quantities $\boldsymbol{B}$, $\boldsymbol{\upsilon}$, $\boldsymbol{\varepsilon}$ appearing in equation (161) refer to the same point in spacetime. On the contrary, in equation (159) the velocity $\boldsymbol{u}$ of the material particle does not refer to point A, where the electromagnetic field is manifested. Equation (161) also holds for the case where the material particle is in arbitrary motion, as we shall see in a later paragraph.

4.3 The density of energy and momentum in the surrounding spacetime of a material point particle

In the case of the rest mass we cannot apply Gauss's law in order to calculate the energy density D in the surrounding spacetime of the material particle. Because of this we will develop a completely different proving procedure. We initially calculate the energy density D' in the inertial reference frame S' in which the material particle is at rest. At point A of figure 7 the energy density D' due to the selfvariations is

$$D' = c^2 \frac{m_0 \left(t' - \dfrac{r'}{c} \right) - m_0 \left(t' - \dfrac{r' + dr'}{c} \right)}{4\pi r'^2 dr'}. \tag{162}$$

From equation (112) and for a specific time t' we have that

$$dw' = -\frac{dr'}{c}$$

and equation (162) becomes

$$D' = c^2 \frac{\dfrac{dm_0}{dr'}}{4\pi r'^2} = -c \frac{\dfrac{dm_0}{dw'}}{4\pi r'^2}. \tag{163}$$

We now consider the Lorentz-Einstein transformations for the energy E and the momentum $\boldsymbol{P}$ of the generalized photon:

$$\begin{array}{ll} E = \gamma (E' + u P'_x) & E' = \gamma (E - u P_x) \\ P_x = \gamma \left(P'_x + \dfrac{u}{c^2} E' \right) & P'_x = \gamma \left(P_x - \dfrac{u}{c^2} E \right) \\ P_y = P'_y & P'_y = P_y \\ P_z = P'_z & P'_z = P_z. \end{array} \tag{164}$$

Defining as dV the infinitesimal volume occupied by the generalized photon at point A of figure 6 we have

$$D = \frac{dE}{dV}.$$

Applying the transformations (164) and (150) we get

$$D = \frac{\gamma (dE' + u dP'_x)}{\gamma \left(1 - \dfrac{\boldsymbol{\upsilon} \cdot \boldsymbol{u}}{c^2} \right) dV'}$$

$$D = \frac{dE' + u \dfrac{\upsilon'_x}{c^2} dE'}{\left(1 - \dfrac{\boldsymbol{\upsilon} \cdot \boldsymbol{u}}{c^2} \right) dV'}$$

$$D = \frac{1 + \dfrac{u\upsilon'_x}{c^2}}{1 - \dfrac{\boldsymbol{\upsilon} \cdot \boldsymbol{u}}{c^2}} \frac{dE'}{dV'}$$

$$D = \frac{1 + \dfrac{u\upsilon'_x}{c^2}}{1 - \dfrac{\boldsymbol{\upsilon} \cdot \boldsymbol{u}}{c^2}} D'. \tag{165}$$

From transformations (123) for the velocity we get

$$1 + \frac{u\upsilon'_x}{c^2} = 1 + \frac{u}{c^2} \frac{\upsilon_x - u}{1 - \dfrac{u\upsilon_x}{c^2}} =$$

$$\frac{1 - \dfrac{u^2}{c^2}}{1 - \dfrac{u\upsilon_x}{c^2}} = \frac{1}{\gamma^2 \left(1 - \dfrac{u\upsilon_x}{c^2} \right)}$$

and since $\frac{\boldsymbol{\upsilon} \cdot \boldsymbol{u}}{c^2} = \frac{u}{c} \cos\delta$, we get

$$1 + \frac{u\upsilon'_x}{c^2} = \frac{1}{\gamma^2 \left(1 - \dfrac{\boldsymbol{\upsilon} \cdot \boldsymbol{u}}{c^2} \right)}. \tag{166}$$

Combining equations (165) and (166) we have

$$D = \frac{1}{\gamma^2\left(1-\frac{\boldsymbol{\upsilon}\cdot\boldsymbol{u}}{c^2}\right)^2}D'$$

and with (163) we get

$$D = -c\frac{1}{\gamma^2\left(1-\frac{\boldsymbol{\upsilon}\cdot\boldsymbol{u}}{c^2}\right)^2}\frac{\frac{dm_0}{dw'}}{4\pi r'^2}.$$

Applying transformations (120) and (127) we obtain

$$D = -c\frac{\partial m_0}{\partial w}\frac{1}{4\pi\gamma^3 r^2\left(1-\frac{\boldsymbol{\upsilon}\cdot\boldsymbol{u}}{c^2}\right)^4}. \tag{167}$$

The generalized photon moves with velocity $\boldsymbol{\upsilon}$, so we have

$$\boldsymbol{J} = \boldsymbol{D}\frac{\boldsymbol{\upsilon}}{c^2} \tag{168}$$

for the momentum density $\boldsymbol{J}$ at point A of figure 6.

Factor $\frac{\partial m_0}{\partial w}$, which appears in the equations of this paragraph, corresponds to factor $\frac{\partial q}{\partial w}$ in the equations of the previous paragraph. In figure 6, the rest mass m_0 of the point particle acts on point $A\,(x, y, x, t)$ with the value it had at point E, namely $m_0 = m_0\,(w)$. Therefore, we have

$$\frac{\partial m_0}{\partial t} = \frac{\partial m_0}{\partial w}\frac{\partial w}{\partial t}$$
$$\nabla m_0 = \frac{\partial m_0}{\partial w}\nabla w$$

and with equations (11) and (12), we get

$$\begin{aligned}\frac{\partial m_0}{\partial t} &= \frac{\partial m_0}{\partial w}\frac{1}{1-\frac{\boldsymbol{\upsilon}\cdot\boldsymbol{u}}{c^2}}\\ \nabla m_0 &= -\frac{\partial m_0}{c\partial w}\frac{1}{1-\frac{\boldsymbol{\upsilon}\cdot\boldsymbol{u}}{c^2}}\frac{\boldsymbol{\upsilon}}{c}.\end{aligned} \tag{169}$$

These equations are analogous to equations (151) and (152) for the electric charge.

4.4 The selfvariations are in accordance with the principle of conservation of the electric charge

In figure 6 and for the time interval from $w = t - \frac{r}{c}$ to t, the generalized photons emitted by the material particle are contained within a sphere with centre E and radius r. In order for the conservation of the electric charge to hold, we have to prove the validity of equation:

$$q\left(t-\frac{r}{c}\right) = q\,(t) + \int_V \rho dV = q\,(t) + q_i \tag{170}$$

where V is the volume of the sphere with centre E and radius r, and

$$q_i = \int_V \rho dV \tag{171}$$

is the electric charge, due to the selfvariations, contained within the sphere. From equation (145), we get for the infinitesimal volume dV

$$\begin{aligned}&dV = b^2\left(1-\frac{u}{c}\cos\delta\right)\sin\delta d\delta d\omega c dw\\ &0 \le \delta \le \pi\\ &0 \le \omega < 2\pi\\ &0 \le b \le r\\ &t - \frac{r}{c} \le w \le t.\end{aligned} \tag{172}$$

Combining equations (156) and (129) we get

$$\rho = -\frac{\partial q}{c\partial w}\frac{1}{4\pi\gamma^2 r^2\left(1-\frac{u}{c}\cos\delta\right)^3}. \tag{173}$$

Combining equations (171) and (173) we also get

$$q_i = \int_V \rho dV$$
$$q_i = -\int_0^{\pi}\int_0^{2\pi}\int_{t-\frac{r}{c}}^{t}\frac{\partial q}{c\partial w}\frac{1}{4\pi\gamma^2 b^2\left(1-\frac{u}{c}\cos\delta\right)^3} b^2\left(1-\frac{u}{c}\cos\delta\right)\sin\delta d\delta d\omega c dw$$

$$q_i = -\int_0^{\pi}\int_0^{2\pi}\int_{t-\frac{r}{c}}^{t}\frac{\partial q}{\partial w}\frac{\sin\delta}{4\pi\gamma^2\left(1-\frac{u}{c}\cos\delta\right)^2} d\delta d\omega c dw$$

$$q_i = -\frac{1}{2\gamma^2}\int_0^{\pi}\int_{t-\frac{r}{c}}^{t}\frac{\partial q}{\partial w}\frac{\sin\delta}{\left(1-\frac{u}{c}\cos\delta\right)^2} d\delta dw. \tag{174}$$

We now denote

$$\lambda = 1 - \frac{u}{c}\cos\delta. \tag{175}$$

Thus, we have

$$\frac{c}{u}d\lambda = \sin\delta d\delta \tag{176}$$

$$1 - \frac{u}{c} \le \lambda \le 1 + \frac{u}{c}. \tag{177}$$

So we have

$$\int_0^{\pi}\frac{\sin\delta}{\left(1-\frac{u}{c}\cos\delta\right)^2}d\delta = \int_{1-\frac{u}{c}}^{1+\frac{u}{c}}\frac{\frac{c}{u}d\lambda}{\lambda^2} = -\frac{c}{u}\left[\frac{1}{\lambda}\right]_{1-\frac{u}{c}}^{1+\frac{u}{c}} =$$
$$-\frac{c}{u}\left(\frac{1}{1+\frac{u}{c}} - \frac{1}{1-\frac{u}{c}}\right) = -\frac{c}{u}\frac{-2\frac{u}{c}}{1-\frac{u^2}{c^2}} = \frac{2}{1-\frac{u^2}{c^2}} = 2\gamma^2$$

and equation (174) becomes

$$q_i = -\int_{t-\frac{r}{c}}^{t} \frac{\partial q}{\partial w} dw$$
$$q_i = -\left[q\left(w\right)\right]_{t-\frac{r}{c}}^{t}$$
$$q_i = q\left(t-\frac{r}{c}\right) - q\left(t\right)$$
$$q\left(t\right) + q_i = q\left(t-\frac{r}{c}\right)$$

which is equation (170).

We can also prove the conservation of the electric charge through the equation of continuity

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \boldsymbol{j} = 0. \tag{178}$$

Indeed, taking into account equation (157) we have

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \boldsymbol{j} = \frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \boldsymbol{\upsilon})$$
$$\frac{\partial \rho}{\partial t} + \nabla \cdot \boldsymbol{j} = \frac{\partial \rho}{\partial t} + \boldsymbol{\upsilon} \cdot \nabla \rho + \rho \nabla \cdot \boldsymbol{\upsilon}$$

and with equation (22) we get

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \boldsymbol{j} = \frac{\partial \rho}{\partial t} + \boldsymbol{\upsilon} \cdot \nabla \rho + \frac{2c}{r}\rho.$$

Applying equation (86) of the fundamental mathematical theorem, for $f = \rho$, we get

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \boldsymbol{j} = c\frac{\partial \rho}{\partial r} + \frac{2c}{r}\rho. \tag{179}$$

From equation (173) we have

$$\frac{\partial \rho}{\partial r} = -\frac{2\rho}{r}. \tag{180}$$

Combining equations (179) and (180) we finally get

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \boldsymbol{j} = 0.$$

4.5 The selfvariations are in accordance with the conservation principles of energy and momentum

In figure 6, for the time interval from $w = t - \frac{r}{c}$ to t, the generalized photons emitted by the material particle due to the selfvariation of the rest mass are contained within the sphere with centre E and radius r. In order for the conservation of energy to hold, it is enough to prove the validity of the following equation:

$$c^2 \gamma m_0 \left(t-\frac{r}{c}\right) = c^2 \gamma m_0\left(t\right) + \int_V D dV = c^2 m_0\left(t\right) + E_i \tag{181}$$

where V is the volume of the sphere with centre E and radius r, and

$$E_i = \int_V D dV \tag{182}$$

is the energy due to the selfvariation of the rest mass, which is contained within the sphere. Combining equations (167) and (129) we get

$$D = -c\frac{\partial m_0}{\partial w}\frac{1}{4\pi\gamma^3 r^2\left(1-\frac{u}{c}\cos\delta\right)^4}. \tag{183}$$

Combining equations (182) and (183), and following the notation of equation (172), we get

$$E_i = -c\int_0^{\pi}\int_0^{2\pi}\int_{t-\frac{r}{c}}^{t}\frac{\partial m_0}{\partial w}\frac{1}{4\pi\gamma^3 b^2\left(1-\frac{u}{c}\cos\delta\right)^4} b^2\left(1-\frac{u}{c}\cos\delta\right)\sin\delta d\delta d\omega c dw$$

$$E_i = -\frac{c^2}{4\pi\gamma^3}\int_0^{\pi}\int_0^{2\pi}\int_{t-\frac{r}{c}}^{t}\frac{\partial m_0}{\partial w}\frac{\sin\delta}{\left(1-\frac{u}{c}\cos\delta\right)^3} d\delta d\omega dw$$

$$E_i = -\frac{c^2}{2\gamma^3}\int_0^{\pi}\int_{t-\frac{r}{c}}^{t}\frac{\partial m_0}{\partial w}\frac{\sin\delta}{\left(1-\frac{u}{c}\cos\delta\right)^3} d\delta dw. \tag{184}$$

Using the notation of equations (175), (176), and (177) we have

$$\int_0^{\pi}\frac{\sin\delta}{\left(1-\frac{u}{c}\cos\delta\right)^3}d\delta = \int_{1-\frac{u}{c}}^{1+\frac{u}{c}}\frac{u}{\lambda^3}d\lambda =$$
$$-\frac{c}{2u}\left[\frac{1}{\lambda^2}\right]_{1-\frac{u}{c}}^{1+\frac{u}{c}} = -\frac{c}{2u}\left(\frac{1}{\left(1+\frac{u}{c}\right)^2} - \frac{1}{\left(1-\frac{u}{c}\right)^2}\right) =$$
$$-\frac{c}{2u}\frac{-4\frac{u}{c}}{\left(1-\frac{u^2}{c^2}\right)^2} = \frac{2}{\left(1-\frac{u^2}{c^2}\right)^2} = 2\gamma^4.$$

Now (184) becomes

$$E_i = -c^2\gamma\int_{t-\frac{r}{c}}^{t}\frac{\partial m_0}{\partial w}dw$$
$$E_i = -c^2\gamma\left[m_0\right]_{t-\frac{r}{c}}^{t}$$
$$E_i = -c^2\gamma m_0\left(t\right) + c^2\gamma m_0\left(t-\frac{r}{c}\right)$$
$$c^2\gamma m_0\left(t-\frac{r}{c}\right) = c^2\gamma m_0\left(t\right) + E_i$$

which is equation (181).

The conservation of energy can also be proven using the continuity equation

$$\frac{\partial D}{c^2\partial t} + \nabla \cdot \boldsymbol{j} = 0. \tag{185}$$

Indeed, if we take into account equation (168) we obtain

$$\frac{\partial D}{c^2\partial t}+\nabla\cdot\boldsymbol{j}=\frac{\partial D}{c^2\partial t}+\nabla\cdot\left(D\frac{\boldsymbol{\upsilon}}{c^2}\right)$$

$$\frac{\partial D}{c^2\partial t}+\nabla\cdot\boldsymbol{j}=\frac{\partial D}{c^2\partial t}+\frac{\boldsymbol{\upsilon}}{c^2}\nabla D+\frac{D}{c^2}\nabla\cdot\boldsymbol{\upsilon}$$

and with equation (22) we have

$$\frac{\partial D}{c^2\partial t}+\nabla\cdot\boldsymbol{j}=\frac{\partial D}{c^2\partial t}+\frac{\boldsymbol{\upsilon}}{c^2}\nabla D+\frac{D}{c^2}\nabla\frac{2c}{r}.$$

Using equation (86) of the fundamental mathematical theorem for $f=D$, we get

$$\frac{\partial D}{c^2\partial t}+\nabla\cdot\boldsymbol{j}=c\frac{\partial D}{c^2\partial r}+\frac{D}{c^2}\frac{2c}{r}. \tag{186}$$

From equation (183) we have

$$\frac{\partial D}{\partial r}=-\frac{2D}{r}. \tag{187}$$

Combining equations (186) and (187) we get

$$\frac{\partial D}{c^2\partial t}+\nabla\cdot\boldsymbol{j}=0.$$

In order to prove the conservation of momentum, it suffices to prove the corresponding of equation (181), that is, it is enough to prove equation

$$\gamma m_0\left(t-\frac{r}{c}\right)\boldsymbol{u}=\gamma m_0(t)\,\boldsymbol{u}+\int_V\boldsymbol{J}dV=\gamma m_0(t)\,\boldsymbol{u}+\boldsymbol{P}_i \tag{188}$$

where

$$\boldsymbol{P}_i=\int_V\boldsymbol{J}dV \tag{189}$$

is the momentum due to the selfvariation of the rest mass, contained within the sphere of centre E and radius r. Combining equations (189) and (168) we obtain

$$\boldsymbol{P}_i=\int_V D\frac{\boldsymbol{\upsilon}}{c^2}dV. \tag{190}$$

We first work on the x-axis:

$$P_{ix}=\int_V D\frac{\upsilon_x}{c^2}dV.$$

Using equation (121) we get

$$P_{ix}=\int_V D\frac{\cos\delta}{c}dV$$

and with equations (183) and (172) we get

$$P_{ix}=-\int_0^{\pi}\int_0^{2\pi}\int_{t-\frac{r}{c}}^{t}\frac{\partial m_0}{\partial w}\frac{\cos\delta}{4\pi\gamma^3b^2\left(1-\frac{u}{c}\cos\delta\right)^4}b^2\left(1-\frac{u}{c}\cos\delta\right)\sin\delta d\delta d\omega cdw$$

$$P_{ix}=-\int_0^{\pi}\int_0^{2\pi}\int_{t-\frac{r}{c}}^{t}\frac{\partial m_0}{\partial w}\frac{\cos\delta\sin\delta}{4\pi\gamma^3b^2\left(1-\frac{u}{c}\cos\delta\right)^3}d\delta d\omega cdw$$

$$P_{ix}=-\frac{c}{4\pi\gamma^3}\int_0^{\pi}\int_0^{2\pi}\int_{t-\frac{r}{c}}^{t}\frac{\partial m_0}{\partial w}\frac{\cos\delta\sin\delta}{\left(1-\frac{u}{c}\cos\delta\right)^3}d\delta d\omega dw$$

$$P_{ix}=-\frac{c}{2\gamma^3}\int_0^{\pi}\int_{t-\frac{r}{c}}^{t}\frac{\partial m_0}{\partial w}\frac{\cos\delta\sin\delta}{\left(1-\frac{u}{c}\cos\delta\right)^3}d\delta dw. \tag{191}$$

Using the notation appearing in equations (175), (176), and (177) we have

$$\int_0^{\pi}\frac{\cos\delta\sin\delta}{\left(1-\frac{u}{c}\cos\delta\right)^3}d\delta=\int_{1-\frac{u}{c}}^{1+\frac{u}{c}}\frac{c^2}{u^2}\frac{1-\lambda}{\lambda^3}d\lambda=$$

$$\frac{c^2}{u^2}\int_{1-\frac{u}{c}}^{1+\frac{u}{c}}\left(\frac{1}{\lambda^3}-\frac{1}{\lambda^2}\right)d\lambda=\frac{c^2}{u^2}\left(-\frac{1}{2}\left[\frac{1}{\lambda^2}\right]_{1-\frac{u}{c}}^{1+\frac{u}{c}}+\left[\frac{1}{\lambda}\right]_{1-\frac{u}{c}}^{1+\frac{u}{c}}\right)=$$

$$\frac{c^2}{u^2}\left(-\frac{1}{2}\frac{\left(1+\frac{u}{c}\right)^2-\left(1-\frac{u}{c}\right)^2}{\left(1-\frac{u^2}{c^2}\right)^2}+\frac{-2\frac{u}{c}}{1-\frac{u^2}{c^2}}\right)=$$

$$\frac{c^2}{u^2}\left(-\frac{1}{2}\frac{-4\frac{u}{c}}{\left(1-\frac{u^2}{c^2}\right)^2}-\frac{2\frac{u}{c}}{1-\frac{u^2}{c^2}}\right)=$$

$$\frac{2c}{u}\left(\frac{1}{\left(1-\frac{u^2}{c^2}\right)^2}-\frac{1}{1-\frac{u^2}{c^2}}\right)=$$

$$\frac{\frac{2c}{u}}{\left(1-\frac{u^2}{c^2}\right)^2}\left(1-1+\frac{u^2}{c^2}\right)=\frac{\frac{2u}{c}}{\left(1-\frac{u^2}{c^2}\right)^2}=2\gamma^4\frac{u}{c}$$

and equation (181) becomes

$$P_{ix}=-u\gamma\int_{t-\frac{r}{c}}^{t}\frac{\partial m_0}{\partial w}dw=-u\gamma\left[m_0\right]_{t-\frac{r}{c}}^{t}$$

$$P_{ix}=u\gamma m_0\left(t-\frac{r}{c}\right)-u\gamma m_0(t). \tag{192}$$

Similarly for the y-axis we get

$$P_{iy}=-\int_0^{\pi}\int_0^{2\pi}\int_{t-\frac{r}{c}}^{t}\frac{\partial m_0}{\partial w}\frac{\sin\delta}{\left(1-\frac{u}{c}\cos\delta\right)^3}\upsilon_y d\delta d\omega dw$$

and with equation (121)

$$\frac{v_y}{c} = \sin\delta\cos\omega$$

we get

$$P_{iy} = -\int_0^{\pi}\int_0^{2\pi}\int_{t-\frac{r}{c}}^{t}\frac{\partial m_0}{\partial w}\frac{c\sin^2\delta\cos\omega}{\left(1-\frac{u}{c}\cos\delta\right)^3}d\delta d\omega dw. \quad (193)$$

The presence of factor $\cos\omega$ causes integral (193) to vanish, and we have

$$P_{iy} = 0. \quad (194)$$

We can similarly prove that

$$P_{iz} = 0. \quad (195)$$

Given that

$$\boldsymbol{u} = \begin{bmatrix} u \\ 0 \\ 0 \end{bmatrix}$$

equations (192), (194) and (195) can be written as

$$\boldsymbol{P}_i = \boldsymbol{u}\gamma m_0\left(t-\frac{r}{c}\right) - \boldsymbol{u}\gamma m_0(t)$$

which is equation (188).

From equation (181) we get

$$E_i = \int_V DdV = c^2\gamma\left(m_0\left(t-\frac{r}{c}\right) - m_0(t)\right). \quad (196)$$

From equation (188) we also have

$$\boldsymbol{P}_i = \int_V \boldsymbol{J}dV = \boldsymbol{u}\gamma\left(m_0\left(t-\frac{r}{c}\right) - m_0(t)\right). \quad (197)$$

Combining equations (196) and (197) we get

$$\boldsymbol{P}_i = E_i\frac{\boldsymbol{u}}{c^2} \quad (198)$$

and

$$\int_V \boldsymbol{J}dV = \frac{\boldsymbol{u}}{c^2}\int_V DdV. \quad (199)$$

Equations (198) and (199) hold for every volume V, i.e. for every radius r of the sphere with centre E and radius r of figure 6. Therefore, they also hold for $r = 0$, that is, on the material particle at time w. Hence, the total energy E_s and the total momentum $\boldsymbol{P}_s$ emitted by the material particle at time w in all directions, are connected through the relation

$$\boldsymbol{P}_s = E_s\frac{\boldsymbol{u}}{c^2} \quad (200)$$

where $\boldsymbol{u} = \boldsymbol{u}(w)$. This equation has fundamental consequences for the material particle, and we shall encounter them as our study continues.

4.6 The electromagnetic field in the macrocosm. The electromagnetic potential of the selfvariations

Using the symbols at point $A(x, y, z, t)$ of figure 4, the scalar potential V and the vector potential $\boldsymbol{A}$ of the selfvariations are given by the following equations:

$$V = \frac{q\left(1-\frac{u^2}{c^2}\right)}{4\pi\varepsilon_0 r\left(1-\frac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}\right)^2} + \frac{q(\boldsymbol{v}\cdot\boldsymbol{\alpha})}{4\pi\varepsilon_0 c^3\left(1-\frac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}\right)^2} \quad (201)$$

$$\boldsymbol{A} = V\frac{\boldsymbol{v}}{c^2}. \quad (202)$$

The intensity $\boldsymbol{\varepsilon}$ of the electric field, and the intensity $\boldsymbol{B}$ of the magnetic field arising from these two potentials, are given by

$$\boldsymbol{\varepsilon} = \frac{q\left(1-\frac{u^2}{c^2}\right)}{4\pi\varepsilon_0 r^2\left(1-\frac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}\right)^3}\left(\frac{\boldsymbol{v}}{c}-\frac{\boldsymbol{u}}{c}\right) + \\ + \frac{q}{4\pi\varepsilon_0 r\left(1-\frac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}\right)^2}\left[\frac{\left(\frac{\boldsymbol{v}}{c}\boldsymbol{\alpha}\right)}{1-\frac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}}\left(\frac{\boldsymbol{v}}{c}-\frac{\boldsymbol{u}}{c}\right)-\boldsymbol{\alpha}\right] \quad (203)$$

$$\mathrm{B} = \frac{q\left(1-\frac{u^2}{c^2}\right)}{4\pi\varepsilon_0 r^2\left(1-\frac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}\right)^3}\frac{\boldsymbol{u}}{c}\times\frac{\boldsymbol{v}}{c} + \\ + \frac{q}{4\pi\varepsilon_0 r\left(1-\frac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}\right)}\left[\frac{\left(\frac{\boldsymbol{v}}{c}\boldsymbol{\alpha}\right)}{1-\frac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}}\left(\frac{\boldsymbol{u}}{c}\times\frac{\boldsymbol{v}}{c}\right)-\frac{\boldsymbol{v}}{c}\times\boldsymbol{\alpha}\right] \quad (204)$$

where $\boldsymbol{u} = \boldsymbol{u}(w)$ is the velocity, and $\boldsymbol{\alpha} = \boldsymbol{\alpha}(w)$ is the acceleration of the material particle. Furthermore, the density of electric charge at point A is

$$\rho = -\frac{\partial q}{\partial w}\frac{1}{4\pi\gamma^2 r^2\left(1-\frac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}\right)^3} \quad (205)$$

exactly as given by equation (156).

In equations (203) and (204) we recognize the electromagnetic field as we know it experimentally, but also as predicted by the Lienard-Wiechert potentials. However, the electromagnetic potentials of the selfvariations have a fundamental characteristic that is not shared by the Lienard-Wiechert potentials. Namely, they split into two individual couples of potentials

$$V_u = \frac{q\left(1-\frac{u^2}{c^2}\right)}{4\pi\varepsilon_0 r\left(1-\frac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}\right)^2} \quad (206)$$

$$\boldsymbol{A}_u = V_u\frac{\boldsymbol{v}}{c^2}$$

and

$$V_\alpha = \frac{q\,(\boldsymbol{\upsilon}\cdot\boldsymbol{\alpha})}{4\pi\varepsilon_0 c^3\left(1-\dfrac{\boldsymbol{\upsilon}\cdot\boldsymbol{u}}{c^2}\right)^2}$$
$$\boldsymbol{A}_\alpha = V_\alpha\frac{\boldsymbol{\upsilon}}{c^2}. \tag{207}$$

The (206) potentials express the electromagnetic field

$$\boldsymbol{\varepsilon}_u = \frac{q\left(1-\dfrac{u^2}{c^2}\right)}{4\pi\varepsilon_0 r^2\left(1-\dfrac{\boldsymbol{\upsilon}\cdot\boldsymbol{u}}{c^2}\right)^3}\left(\frac{\boldsymbol{\upsilon}}{c}-\frac{\boldsymbol{u}}{c}\right)$$
$$\mathrm{B}_u = \frac{q\left(1-\dfrac{u^2}{c^2}\right)}{4\pi\varepsilon_0 r^2\left(1-\dfrac{\boldsymbol{\upsilon}\cdot\boldsymbol{u}}{c^2}\right)^3}\frac{\boldsymbol{u}}{c}\times\frac{\boldsymbol{\upsilon}}{c}. \tag{208}$$

that accompanies the material particle in its motion. The (207) potentials express the electromagnetic radiation

$$\boldsymbol{\varepsilon}_\alpha = \frac{q}{4\pi\varepsilon_0 c^2 r\left(1-\dfrac{\boldsymbol{\upsilon}\cdot\boldsymbol{u}}{c^2}\right)^2}\left[\frac{\left(\dfrac{\boldsymbol{\upsilon}}{c}\boldsymbol{\alpha}\right)}{1-\dfrac{\boldsymbol{\upsilon}\cdot\boldsymbol{u}}{c^2}}\left(\frac{\boldsymbol{\upsilon}}{c}-\frac{\boldsymbol{u}}{c}\right)-\boldsymbol{\alpha}\right]$$
$$\mathrm{B}_\alpha = \frac{q}{4\pi\varepsilon_0 r\left(1-\dfrac{\boldsymbol{\upsilon}\cdot\boldsymbol{u}}{c^2}\right)}\left[\frac{\left(\dfrac{\boldsymbol{\upsilon}}{c}\boldsymbol{\alpha}\right)}{1-\dfrac{\boldsymbol{\upsilon}\cdot\boldsymbol{u}}{c^2}}\left(\frac{\boldsymbol{u}}{c}\times\frac{\boldsymbol{\upsilon}}{c}\right)-\frac{\boldsymbol{\upsilon}}{c}\times\boldsymbol{\alpha}\right]. \tag{209}$$

The (207) potential of the electromagnetic radiation does not depend on the distance r, while it vanishes for $\boldsymbol{\upsilon}\cdot\boldsymbol{\alpha}=0$. Furthermore, for each couple of the electromagnetic field we can easily prove that equation (161) holds

$$\mathrm{B}_u = \frac{\boldsymbol{\upsilon}}{c^2}\times\boldsymbol{\varepsilon}_u \tag{210}$$

$$\mathrm{B}_\alpha = \frac{\boldsymbol{\upsilon}}{c^2}\times\boldsymbol{\varepsilon}_\alpha. \tag{211}$$

We remind the reader that the electromagnetic field can be calculated from the electromagnetic potentials via equations

$$\boldsymbol{\varepsilon} = -\nabla V - \frac{\partial \mathrm{A}}{\partial t} \tag{212}$$

$$\mathrm{B} = \nabla\times\mathrm{A} \tag{213}$$

where $\nabla V = \begin{bmatrix}\frac{\partial V}{\partial x}\\ \frac{\partial V}{\partial y}\\ \frac{\partial V}{\partial z}\end{bmatrix}$, and $\nabla\times\boldsymbol{A} = \operatorname{curl}\boldsymbol{A}$.

We shall now prove equation

$$\boldsymbol{\varepsilon}_\alpha = -\nabla V_\alpha - \frac{\partial \mathrm{A}_\alpha}{\partial t} \tag{214}$$

and the general equations (203) and (204) can be proven similarly. We shall make use of the equations of paragraph 2.7. From the (207) potentials we obtain

$$\boldsymbol{\varepsilon}_\alpha = -\left[\frac{q\,(\boldsymbol{\upsilon}\cdot\boldsymbol{\alpha})}{4\pi\varepsilon_0 c^3\left(1-\dfrac{\boldsymbol{\upsilon}\cdot\boldsymbol{u}}{c^2}\right)^2}\right]-\frac{\partial}{\partial t}\left[\frac{q\,(\boldsymbol{\upsilon}\cdot\boldsymbol{\alpha})}{4\pi\varepsilon_0 c^5\left(1-\dfrac{\boldsymbol{\upsilon}\cdot\boldsymbol{u}}{c^2}\right)^2}\boldsymbol{\upsilon}\right]$$
$$\boldsymbol{\varepsilon}_\alpha = -\nabla\left[\frac{q\,(\boldsymbol{\upsilon}\cdot\boldsymbol{\alpha})}{4\pi\varepsilon_0 c^3\left(1-\dfrac{\boldsymbol{\upsilon}\cdot\boldsymbol{u}}{c^2}\right)^2}\right]-\boldsymbol{\upsilon}\frac{\partial}{\partial t}\left[\frac{q\,(\boldsymbol{\upsilon}\cdot\boldsymbol{\alpha})}{4\pi\varepsilon_0 c^5\left(1-\dfrac{\boldsymbol{\upsilon}\cdot\boldsymbol{u}}{c^2}\right)^2}\right]$$
$$-\frac{q\,(\boldsymbol{\upsilon}\cdot\boldsymbol{\alpha})}{4\pi\varepsilon_0 c^5\left(1-\dfrac{\boldsymbol{\upsilon}\cdot\boldsymbol{u}}{c^2}\right)^2}\frac{\partial\boldsymbol{\upsilon}}{\partial t}$$

$$\boldsymbol{\varepsilon}_\alpha = -\frac{(\boldsymbol{\upsilon}\cdot\boldsymbol{\alpha})}{4\pi\varepsilon_0\left(1-\dfrac{\boldsymbol{\upsilon}\cdot\boldsymbol{u}}{c^2}\right)^2}\left(\nabla q+\frac{\partial q}{c^2\partial t}\boldsymbol{\upsilon}\right)-$$
$$-\frac{q}{4\pi\varepsilon_0\left(1-\dfrac{\boldsymbol{\upsilon}\cdot\boldsymbol{u}}{c^2}\right)^2}\left[\nabla(\boldsymbol{\upsilon}\cdot\boldsymbol{\alpha})+\frac{\partial(\boldsymbol{\upsilon}\cdot\boldsymbol{\alpha})}{c^2\partial t}\boldsymbol{\upsilon}\right]-$$
$$-\frac{2q\,(\boldsymbol{\upsilon}\cdot\boldsymbol{\alpha})}{4\pi\varepsilon_0 c^3\left(1-\dfrac{\boldsymbol{\upsilon}\cdot\boldsymbol{u}}{c^2}\right)^3}\left[\nabla\left(\frac{\boldsymbol{\upsilon}\cdot\boldsymbol{u}}{c^2}\right)+\frac{\partial(\boldsymbol{\upsilon}\cdot\boldsymbol{u})}{c^4\partial t}\boldsymbol{\upsilon}\right]- \tag{215}$$
$$-\frac{q\,(\boldsymbol{\upsilon}\cdot\boldsymbol{\alpha})}{4\pi\varepsilon_0 c^5\left(1-\dfrac{\boldsymbol{\upsilon}\cdot\boldsymbol{u}}{c^2}\right)^2}\frac{\partial\boldsymbol{\upsilon}}{\partial t}.$$

Combining equations (151) and (152) we get

$$\nabla q+\frac{\partial q}{c^2\partial t}\boldsymbol{\upsilon}=0. \tag{216}$$

Combining equations (98) and (99) we get

$$\nabla(\boldsymbol{\upsilon}\cdot\boldsymbol{\alpha})+\frac{\partial(\boldsymbol{\upsilon}\cdot\boldsymbol{\alpha})}{c^2\partial t}=\frac{c}{r}\boldsymbol{\alpha}-\frac{(\boldsymbol{\upsilon}\cdot\boldsymbol{\alpha})}{cr}\boldsymbol{\upsilon}. \tag{217}$$

Combining equations (96) and (99) we get

$$\nabla(\boldsymbol{\upsilon}\cdot\boldsymbol{u})+\frac{\partial(\boldsymbol{\upsilon}\cdot\boldsymbol{u})}{c^2\partial t}\boldsymbol{\upsilon}=-\frac{c}{r}\left(\frac{(\boldsymbol{\upsilon}\cdot\boldsymbol{u})}{c^2}\boldsymbol{\upsilon}-u\right). \tag{218}$$

We substitute equations (216), (217) and (218) into equation (215) and we obtain

$$\boldsymbol{\varepsilon}_\alpha = -\frac{q}{4\pi\varepsilon_0 c^3\left(1-\frac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}\right)^2}\left[\frac{c}{r}\boldsymbol{\alpha}-\frac{(\boldsymbol{v}\cdot\boldsymbol{\alpha})}{cr}\boldsymbol{v}\right]+$$

$$+\frac{2q(\boldsymbol{v}\cdot\boldsymbol{\alpha})}{4\pi\varepsilon_0 c^5\left(1-\frac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}\right)^3}\frac{c}{r}\left[\frac{(\boldsymbol{v}\cdot\boldsymbol{u})}{c^2}\boldsymbol{v}-\boldsymbol{u}\right]-$$

$$-\frac{q(\boldsymbol{v}\cdot\boldsymbol{\alpha})}{4\pi\varepsilon_0 c^4\left(1-\frac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}\right)^3}\left[\frac{(\boldsymbol{v}\cdot\boldsymbol{u})}{c^2}\boldsymbol{v}-\boldsymbol{u}\right]$$

$$\boldsymbol{\varepsilon}_\alpha = -\frac{q}{4\pi\varepsilon_0 c^3\left(1-\frac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}\right)^2}\left[\frac{c}{r}\boldsymbol{\alpha}-\frac{(\boldsymbol{v}\cdot\boldsymbol{\alpha})}{cr}\boldsymbol{v}\right]+$$

$$+\frac{q(\boldsymbol{v}\cdot\boldsymbol{\alpha})}{4\pi\varepsilon_0 c^4 r\left(1-\frac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}\right)^3}\left[\frac{(\boldsymbol{v}\cdot\boldsymbol{u})}{c^2}\boldsymbol{v}-\boldsymbol{u}\right]$$

$$\boldsymbol{\varepsilon}_\alpha = \frac{q}{4\pi\varepsilon_0 c^2 r\left(1-\frac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}\right)^2}\cdot$$

$$\left[-\boldsymbol{\alpha}+\frac{(\boldsymbol{v}\cdot\boldsymbol{\alpha})}{c^2}\boldsymbol{v}+\frac{(\boldsymbol{v}\cdot\boldsymbol{u})(\boldsymbol{v}\cdot\boldsymbol{\alpha})}{c^4\left(1-\frac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}\right)}\boldsymbol{v}-\frac{(\boldsymbol{v}\cdot\boldsymbol{\alpha})}{c^2\left(1-\frac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}\right)}\boldsymbol{u}\right]$$

$$\boldsymbol{\varepsilon}_\alpha = \frac{q}{4\pi\varepsilon_0 c^2 r\left(1-\frac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}\right)^2}\cdot$$

$$\left[-\boldsymbol{\alpha}+\frac{(\boldsymbol{v}\cdot\boldsymbol{\alpha})}{c^2\left(1-\frac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}\right)}\left(\left(1-\frac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}\right)\boldsymbol{v}+\frac{(\boldsymbol{v}\cdot\boldsymbol{u})}{c^2}\boldsymbol{v}-\boldsymbol{u}\right)\right]$$

$$\boldsymbol{\varepsilon}_\alpha = \frac{q}{4\pi\varepsilon_0 c^2 r\left(1-\frac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}\right)^2}\left[-\boldsymbol{\alpha}+\frac{(\boldsymbol{v}\cdot\boldsymbol{\alpha})}{c^2\left(1-\frac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}\right)}(\boldsymbol{v}-u)\right]$$

which is equation (209) for the electric field $\boldsymbol{\varepsilon}_\alpha$.

In order to prove equations (208) we also need equations

$$\nabla\left(u^2\right)+\frac{\partial\left(u^2\right)}{c^2\partial t}\boldsymbol{v}=0 \tag{219}$$

$$\nabla r+\frac{\partial r}{c^2\partial t}\boldsymbol{v}=\frac{\boldsymbol{v}}{c}. \tag{220}$$

We can prove equation (219) as follows

$$\nabla\left(u^2\right)+\frac{\partial\left(u^2\right)}{c^2\partial t}\boldsymbol{v}=\frac{\partial u^2}{\partial w}\nabla w+\frac{\partial u^2}{c^2\partial w}\frac{\partial w}{\partial t}\boldsymbol{v}$$
$$=\frac{\partial u^2}{\partial w}\left(\nabla w+\frac{\partial w}{c^2\partial t}\cdot\boldsymbol{v}\right)=0.$$

This results immediately from the combination of equations (11) and (12). Equation (220) results from the combination of equations (9) and (14).

In order to prove equation (205), we denote

$$\boldsymbol{f}=\frac{1-\frac{u^2}{c^2}}{4\pi\varepsilon_0 r^2\left(1-\frac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}\right)^3}\left(\frac{\boldsymbol{v}}{c}-\frac{\boldsymbol{u}}{c}\right)+$$
$$+\frac{1}{4\pi\varepsilon_0 r\left(1-\frac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}\right)^2}\left[\frac{\left(\frac{v}{c}\boldsymbol{\alpha}\right)}{1-\frac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}}\left(\frac{\boldsymbol{v}}{c}-\frac{\boldsymbol{u}}{c}\right)-\boldsymbol{\alpha}\right] \tag{221}$$

and

$$\boldsymbol{g}=\frac{1-\frac{u^2}{c^2}}{4\pi\varepsilon_0 r\left(1-\frac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}\right)^3}\left(\frac{\boldsymbol{u}}{c}\times\frac{\boldsymbol{v}}{c}\right)+$$
$$+\frac{1}{4\pi\varepsilon_0 r\left(1-\frac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}\right)^2}\left[\frac{\frac{\boldsymbol{v}}{c}\boldsymbol{\alpha}}{1-\frac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}}\left(\frac{\boldsymbol{u}}{c}\times\frac{\boldsymbol{v}}{c}\right)-\frac{\boldsymbol{v}}{c}\times\boldsymbol{\alpha}\right]. \tag{222}$$

Using the notation of equations (221) and (222), and from equations (208) and (209) we obtain

$$\boldsymbol{\varepsilon}=\boldsymbol{\varepsilon}_u+\boldsymbol{\varepsilon}_\alpha=q\boldsymbol{f} \tag{223}$$

$$\boldsymbol{B}=\boldsymbol{B}_u+\boldsymbol{B}_\alpha=q\boldsymbol{g}. \tag{224}$$

From Gauss's law we have

$$\rho=\varepsilon_0\nabla\cdot\boldsymbol{\varepsilon}$$

and using equation (223) we have

$$\rho=\varepsilon_0\nabla\cdot(q\boldsymbol{f})$$
$$\rho=\varepsilon_0 q\nabla\cdot\boldsymbol{f}+\varepsilon_0\boldsymbol{f}\cdot\nabla q. \tag{225}$$

From classical electromagnetism we know that

$$\nabla\cdot\boldsymbol{f}=0.$$

Hence, equation (225) becomes

$$\rho=\varepsilon_0\boldsymbol{f}\cdot\nabla q.$$

Using equation (216) we obtain

$$\rho=-\varepsilon_0\frac{\partial q}{c^2\partial t}\boldsymbol{v}\cdot\boldsymbol{f}. \tag{226}$$

From equation (221) we see that

$$\boldsymbol{v}\cdot\boldsymbol{f}=\frac{1-\frac{u^2}{c^2}}{r^2\left(1-\frac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}\right)^3}\left(\frac{c^2}{c}-\frac{\boldsymbol{v}\cdot\boldsymbol{u}}{c}\right)+$$
$$+\frac{1}{4\pi\varepsilon_0 r\left(1-\frac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}\right)^2}\left[\frac{\left(\frac{\boldsymbol{v}}{c}\boldsymbol{u}\right)}{1-\frac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}}\left(\frac{c^2}{c}-\frac{\boldsymbol{v}\cdot\boldsymbol{u}}{c}\right)-\boldsymbol{v}\cdot\boldsymbol{\alpha}\right]$$

$$\boldsymbol{\upsilon}\cdot\boldsymbol{f}=\frac{c\left(1-\dfrac{u^2}{c^2}\right)}{4\pi\varepsilon_0 r\left(1-\dfrac{\boldsymbol{\upsilon}\cdot\boldsymbol{u}}{c^2}\right)^2}+$$

$$+\frac{1}{r\left(1-\dfrac{\boldsymbol{\upsilon}\cdot\boldsymbol{u}}{c^2}\right)}\left[\frac{(\boldsymbol{\upsilon}\cdot\boldsymbol{\alpha})}{1-\dfrac{\boldsymbol{\upsilon}\cdot\boldsymbol{u}}{c^2}}\left(1-\frac{\boldsymbol{\upsilon}\cdot\boldsymbol{u}}{c^2}\right)-\boldsymbol{\upsilon}\cdot\boldsymbol{\alpha}\right]$$

$$\boldsymbol{\upsilon}\cdot\boldsymbol{f}=\frac{c\left(1-\dfrac{u^2}{c^2}\right)}{r^2\left(1-\dfrac{\boldsymbol{\upsilon}\cdot\boldsymbol{u}}{c^2}\right)^2}+0. \tag{227}$$

Combining equations (226) and (227) we get

$$\rho=-\varepsilon_0\frac{1-\dfrac{u^2}{c^2}}{4\pi\varepsilon_0\left(1-\dfrac{\boldsymbol{\upsilon}\cdot\boldsymbol{u}}{c^2}\right)^2}\frac{\partial q}{\partial t}$$

and with equation (151) we finally obtain

$$\rho=-\frac{\partial q}{\partial w}\frac{1-\dfrac{u^2}{c^2}}{4\pi r^2\left(1-\dfrac{\boldsymbol{\upsilon}\cdot\boldsymbol{u}}{c^2}\right)^3}$$

which is equation (205), since

$$\frac{1}{\gamma^2}=1-\frac{u^2}{c^2}.$$

Similarly we can prove equation

$$\nabla\cdot\boldsymbol{B}=0. \tag{228}$$

From equation (224) we have that

$$\nabla\cdot\boldsymbol{B}=\nabla\cdot(q\boldsymbol{g})$$

$$\nabla\cdot\boldsymbol{B}=q\nabla\cdot\boldsymbol{g}+g\cdot\nabla q. \tag{229}$$

From classical electromagnetism we know that

$$\nabla\cdot\boldsymbol{g}=0.$$

Thus, equation (229) becomes

$$\nabla\cdot\boldsymbol{B}=\boldsymbol{g}\cdot\nabla q$$

and with equation (216) we obtain

$$\nabla\cdot\boldsymbol{B}=-\frac{\partial q}{c^2\partial t}\boldsymbol{\upsilon}\cdot\boldsymbol{g}. \tag{230}$$

From equation (222) it immediately can be seen that

$$\boldsymbol{\upsilon}\cdot\boldsymbol{g}=0$$

and from equation (230) we also obtain

$$\nabla\cdot\boldsymbol{B}=0.$$

Combining equations (230) and (224) we have that

$$\nabla\cdot\boldsymbol{B}=-\frac{\partial q}{c^2 q\partial t}\boldsymbol{\upsilon}\cdot\boldsymbol{B}. \tag{231}$$

From equation (231) it follows that

$$\nabla\cdot\boldsymbol{B}=0$$

if and only if

$$\boldsymbol{\upsilon}\cdot\boldsymbol{B}=0.$$

From equations (210) and (211) we get

$$\boldsymbol{B}=\frac{\boldsymbol{\upsilon}}{c^2}\times\boldsymbol{\varepsilon}. \tag{232}$$

Therefore, it holds that

$$\boldsymbol{\upsilon}\cdot\boldsymbol{B}=0$$

or equivalently

$$\nabla\cdot\boldsymbol{B}=0.$$

Combining equations (226) and (223) we get

$$\rho=-\varepsilon_0\frac{\partial q}{c^2 q\partial t}\boldsymbol{\upsilon}\cdot\boldsymbol{\varepsilon}. \tag{233}$$

From equation (233) it follows that

$$\rho=0$$

if and only if

$$\boldsymbol{\upsilon}\cdot\boldsymbol{\varepsilon}=0.$$

From equation (209) for the electric field $\boldsymbol{\varepsilon}_\alpha$, we can immediately deduce that

$$\boldsymbol{\upsilon}\cdot\boldsymbol{\varepsilon}_\alpha=0. \tag{234}$$

Therefore, the electromagnetic radiation does not contribute to the charge density ρ. On the contrary, for the electric field $\boldsymbol{\varepsilon}_u$ that accompanies the material particle, it holds that

$$\boldsymbol{\upsilon}\cdot\boldsymbol{\varepsilon}_u\neq 0$$

as follows from equation (208).

From equation (232) we obtain

$$\boldsymbol{B}=\frac{\boldsymbol{\upsilon}}{c^2}\times\boldsymbol{\varepsilon}$$

$$\boldsymbol{B}^2=\left(\frac{\boldsymbol{\upsilon}}{c^2}\times\boldsymbol{\varepsilon}\right)^2$$

$$\boldsymbol{B}^2=\left(\frac{\boldsymbol{\upsilon}}{c^2}\times\boldsymbol{\varepsilon}\right)\cdot\left(\frac{\boldsymbol{\upsilon}}{c^2}\times\boldsymbol{\varepsilon}\right).$$

After performing the necessary calculations we finally get

$$\boldsymbol{\varepsilon}^2 = c^2\boldsymbol{B}^2 + \left(\frac{\boldsymbol{\upsilon}\cdot\boldsymbol{\varepsilon}}{c}\right)^2. \tag{235}$$

We end this paragraph with an interesting observation. Comparing equations (208) for the electric field $\boldsymbol{\varepsilon}_u$ with equation (65), we conclude that the vectors $\boldsymbol{t}$ and $\boldsymbol{\varepsilon}_u$ are parallel. Then, the "trajectory representation theorem" informs us that the direction of the electric field $\boldsymbol{\varepsilon}_u$ represents the tangential vector $\boldsymbol{t}_p$ of the trajectory C_p of the material particle.

4.7 The energy-momentum tensor of the electromagnetic field at macroscopic scales

The equations of this paragraph as well as of the remaining paragraphs of this paragraph, could be stated differently, so that they also hold for non-inertial reference frames. However, such a formulation does not serve the purposes of the present edition. Therefore, we will formulate the equations for an inertial reference frame, while simultaneously suggesting the way in which the same equations can also be formulated for a non-inertial reference frame.

From the axiomatic foundation of the theory of selfvariations, as stated in paragraph 2.2, we have that

$$dS^2 = 0$$

or, equivalently,

$$g_{ik}dx^i dx^k = 0 \qquad i,k = 0,1,2,3 \tag{236}$$

where

$$\left(x^0, x^1, x^2, x^3\right) = (ct, x, y, z) \tag{237}$$

and g_{ik} are the components of the metric tensor. In equation (236) we use the Einstein summation convention for the indices i and k.

We denote

$$\upsilon^i = \frac{dx^i}{dt} \qquad i = 0,1,2,3 \tag{238}$$

that is,

$$\left(\upsilon^0, \upsilon^1, \upsilon^2, \upsilon^3\right) = \left(c, \upsilon_x, \upsilon_y, \upsilon_z\right). \tag{239}$$

From equation (236) we obtain

$$g_{ik}\frac{dx^1}{dt}\frac{dx^k}{dt} = 0$$

and with equation (238) we get

$$g_{ik}\upsilon^i\upsilon^k = 0 \qquad i,k = 0,1,2,3. \tag{240}$$

Using this notation, all the equations we will formulate also hold for non-inertial reference frames if we replace differentiation with respect to x^k with covariant differentiation with respect to x^k, k=0,1,2,3.

We now denote the four-vector of velocity as

$$\boldsymbol{\upsilon} = \begin{bmatrix} \upsilon^0 \\ \upsilon^1 \\ \upsilon^2 \\ \upsilon^3 \end{bmatrix} = \begin{bmatrix} c \\ \upsilon_x \\ \upsilon_y \\ \upsilon_z \end{bmatrix} \tag{241}$$

and the four-vector of current density as

$$\boldsymbol{j} = \begin{bmatrix} j^0 \\ j^1 \\ j^2 \\ j^3 \end{bmatrix} = \begin{bmatrix} \rho\upsilon^0 \\ \rho\upsilon^1 \\ \rho\upsilon^2 \\ \rho\upsilon^3 \end{bmatrix} = \begin{bmatrix} \rho c \\ \rho\upsilon_x \\ \rho\upsilon_y \\ \rho\upsilon_z \end{bmatrix} \tag{242}$$

as results from equations (156) and (157). Also, according to equations (201) and (202), the four-vector of the electromagnetic potential is

$$\boldsymbol{A} = \begin{bmatrix} A^0 \\ A^1 \\ A^2 \\ A^3 \end{bmatrix} = \begin{bmatrix} \frac{V}{c}\upsilon^0 \\ \frac{V}{c}\upsilon^1 \\ \frac{V}{c}\upsilon^2 \\ \frac{V}{c}\upsilon^3 \end{bmatrix} = \begin{bmatrix} V \\ \frac{V}{c}\upsilon_x \\ \frac{V}{c}\upsilon_y \\ \frac{V}{c}\upsilon_z \end{bmatrix}. \tag{243}$$

Subsequently we will symbolize the differentiation with respect to $\frac{\partial}{\partial x^k}$ with $(,k)$, k=0,1,2,3.

We now consider the tensor of the electromagnetic field

$$T^{\mu\nu} = \frac{I}{4\pi}\left(F^{\mu\alpha}F^{\nu}_{\alpha} - \frac{1}{4}g^{\mu\nu}F_{\alpha\beta}F^{\alpha\beta}\right) \tag{244}$$

where $g^{\mu\nu}$ is the inverse of the matrix $g_{\mu\nu}$, $g_{\mu k}g^{k\nu} = \delta_{\mu\nu}$

$$\delta_{\mu\nu} = \begin{cases} 1 & for \quad \mu = \nu \\ 0 & for \quad \mu \neq \nu \end{cases} \tag{245}$$

and $F^{\mu\nu}$ is the Maxwell stress tensor

$$F^{\mu\nu} = A^{\nu}_{,\mu} - A^{\mu}_{,\nu}. \tag{246}$$

Using this notation and taking into account that in the surrounding spacetime of the material particle there is an electric current $\boldsymbol{j}$, as given by equation (242), the energy-momentum tensor [19–21] of the electromagnetic field is given by the tensor

$$\Phi^{\mu\nu} = T^{\mu\nu} - j^{\mu}A^{\nu}. \tag{247}$$

We now write the tensor $T^{\mu\nu}$ in the form

$$T^{ij} = \begin{bmatrix} W & cS_x & cS_y & cS_z \\ cS_x & & & \\ cS_y & & \sigma_{\alpha\beta} & \\ cS_z & & & \end{bmatrix} \tag{248}$$

$$\boldsymbol{S} = \varepsilon_0 \boldsymbol{\varepsilon} \times \boldsymbol{B} \tag{249}$$

where $\boldsymbol{S}$ is the Poynting vector, and $\boldsymbol{\varepsilon}$ and B are the intensities of the electric and magnetic field, respectively. Taking into account equations (210) and (211), as summarized in equation

$$\boldsymbol{B} = \frac{\boldsymbol{\upsilon}}{c^2} \times \boldsymbol{\varepsilon} \tag{250}$$

equation (249) becomes

$$\boldsymbol{S} = \varepsilon_0 \left(\frac{\boldsymbol{\varepsilon}^2}{c^2} \right) \boldsymbol{\upsilon} - \varepsilon_0 \left(\frac{\boldsymbol{\upsilon} \cdot \boldsymbol{\varepsilon}}{c^2} \right) \boldsymbol{\varepsilon}. \tag{251}$$

The Maxwell stress tensor $\sigma_{\alpha\beta}$ is given by relation

$$\sigma_{\alpha\beta} = \varepsilon_0 \left(-\varepsilon_\alpha \varepsilon_\beta - c^2 B_\alpha B_\beta + W \delta_{\alpha\beta} \right) \tag{252}$$

where $\delta_{\alpha\beta}$ is given by relation (245), and

$$W = \frac{1}{2} \varepsilon_0 \left(\boldsymbol{\varepsilon}^2 + c^2 \boldsymbol{B}^2 \right) \tag{253}$$

$$\varepsilon = \begin{bmatrix} \varepsilon_x \\ \varepsilon_y \\ \varepsilon_z \end{bmatrix} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{bmatrix}$$

$$B = \begin{bmatrix} B_x \\ B_y \\ B_z \end{bmatrix} = \begin{bmatrix} B_1 \\ B_2 \\ B_3 \end{bmatrix}.$$

Combining equations (247) and (248), we arrive at the energy-momentum tensor

$$\Phi^{ij} = \begin{bmatrix} W & cS_x & cS_y & cS_z \\ cS_x & \sigma_{11} & \sigma_{12} & \sigma_{13} \\ cS_y & \sigma_{21} & \sigma_{22} & \sigma_{23} \\ cS_z & \sigma_{31} & \sigma_{32} & \sigma_{33} \end{bmatrix} - \frac{\rho V}{c^2} \begin{bmatrix} c^2 & c\upsilon_x & c\upsilon_y & c\upsilon_z \\ \upsilon_x c & \upsilon_x^2 & \upsilon_x \upsilon_y & \upsilon_x \upsilon_z \\ \upsilon_y c & \upsilon_y \upsilon_x & \upsilon_y^2 & \upsilon_y \upsilon_z \\ \upsilon_z c & \upsilon_z \upsilon_x & \upsilon_z \upsilon_y & \upsilon_z^2 \end{bmatrix}. \tag{254}$$

We shall now prove that the scalar potential, as given by equation (201), satisfies the relation

$$\frac{\partial V}{\partial t} + \boldsymbol{\upsilon} \cdot \nabla V = -\boldsymbol{\upsilon} \cdot \boldsymbol{\varepsilon}. \tag{255}$$

From equation (212) we have that

$$-\boldsymbol{\upsilon} \cdot \boldsymbol{\varepsilon} = -\boldsymbol{\upsilon} \left[-\nabla V - \frac{\partial \boldsymbol{A}}{\partial t} \right]$$

$$-\boldsymbol{\upsilon} \cdot \boldsymbol{\varepsilon} = \boldsymbol{\upsilon} \left(\nabla V + \frac{\partial \boldsymbol{A}}{\partial t} \right).$$

Using equation (202) we have

$$-\boldsymbol{\upsilon} \cdot \boldsymbol{\varepsilon} = \boldsymbol{\upsilon} \left(\nabla V + \frac{\partial V}{c \partial t} \frac{\boldsymbol{\upsilon}}{c} + \frac{V}{c} \frac{\partial}{\partial t} \left(\frac{\boldsymbol{\upsilon}}{c} \right) \right)$$

$$-\boldsymbol{\upsilon} \cdot \boldsymbol{\varepsilon} = \boldsymbol{\upsilon} \cdot \nabla V + \frac{\partial V}{\partial t} + \frac{V}{c} \boldsymbol{\upsilon} \cdot \frac{\partial}{\partial t} \left(\frac{\boldsymbol{\upsilon}}{c} \right)$$

$$-\boldsymbol{\upsilon} \cdot \boldsymbol{\varepsilon} = \boldsymbol{\upsilon} \cdot \nabla V + \frac{\partial V}{\partial t} + \frac{V}{2c} \frac{\partial}{\partial t} \left(\frac{\boldsymbol{\upsilon}^2}{c} \right)$$

$$-\boldsymbol{\upsilon} \cdot \boldsymbol{\varepsilon} = \boldsymbol{\upsilon} \cdot \nabla V + \frac{\partial V}{\partial t}$$

since $\boldsymbol{\upsilon}^2 = c^2$.

We will now prove the conservation of energy and momentum, as expressed by equation

$$\Phi^{ij}_{\ .j} = \frac{\partial \Phi^{ij}}{\partial x^j} = 0. \tag{256}$$

We begin with an observation which allows us to avoid complex calculations. Equation (256) holds in classical electromagnetic theory, i.e. if we ignore the consequences of the selfvariations and consider the electric charge q constant, both in the electromagnetic potential, as well as in the intensity of the electromagnetic field. Furthermore, $\rho = 0$ in equation (254). Therefore, it is enough to prove that in equation (256) the factors resulting from the selfvariation of the electric charge q, also vanish. Certainly, in equation (254) it holds that $\rho \neq 0$, where the charge density ρ is given by equation (205).

The energy density W of the electromagnetic field as given by equation (253), as well as the Poynting vector S, given by equation (251), are proportional to q^2. Therefore, in our calculations we will have to take into consideration the rate of change of the factor q^2. From equations (151) and (152) we have

$$\frac{\partial q^2}{\partial t} = 2q \frac{\partial q}{\partial t} = \frac{2q}{1 - \dfrac{\boldsymbol{\upsilon} \cdot \boldsymbol{u}}{c^2}} \frac{\partial q}{\partial w}$$

$$\nabla q^2 = 2q \nabla q = -\frac{2q}{1 - \dfrac{\boldsymbol{\upsilon} \cdot \boldsymbol{u}}{c^2}} \frac{\partial q}{\partial w} \frac{\boldsymbol{\upsilon}}{c^2}.$$

Thus, we arrive at equations

$$\begin{aligned} \frac{\partial q^2}{\partial t} &= \frac{2}{1 - \dfrac{\boldsymbol{\upsilon} \cdot \boldsymbol{u}}{c^2}} \frac{\partial q}{q \partial w} q^2 = -2\lambda q^2 \\ \frac{\partial q^2}{\partial x} &= 2\lambda \frac{\upsilon_x}{c^2} q^2 \qquad \frac{\partial q^2}{\partial y} = 2\lambda \frac{\upsilon_y}{c^2} q^2 \\ \frac{\partial q^2}{\partial z} &= 2\lambda \frac{\upsilon_z}{c^2} q^2 \\ \lambda &= -\frac{1}{1 - \dfrac{\boldsymbol{\upsilon} \cdot \boldsymbol{u}}{c^2}} \frac{\partial q}{q \partial w}. \end{aligned} \tag{257}$$

From equation (255), and for $i = 0$, we have that

$$\frac{\partial\Phi^{0j}}{\partial x^j} = \frac{\partial\Phi^{00}}{\partial x^0} + \frac{\partial\Phi^{01}}{\partial x^1} + \frac{\partial\Phi^{02}}{\partial x^2} + \frac{\partial\Phi^{03}}{\partial x^3} =$$
$$\frac{\partial}{c\partial t}(W) \frac{\partial}{\partial x}(cS_x) + \frac{\partial}{\partial y}\left(cS_y\right) + \frac{\partial}{\partial z}(cS_z)$$
$$-\frac{1}{c^2}\left[\frac{\partial}{c\partial t}\left(\rho Vc^2\right) + \frac{\partial}{\partial x}(\rho Vc\upsilon_x) + \frac{\partial}{\partial y}\left(\rho Vc\upsilon_y\right) + \frac{\partial}{\partial z}(\rho Vc\upsilon_z)\right]$$

and using relations (257), which we apply on the quantities W, S_x, S_y, S_z, which are proportional to q^2, we get

$$\frac{\partial\Phi^{0j}}{\partial x^j} = -2\lambda W + 2\lambda\upsilon_x S_x + 2\lambda\upsilon_y S_y + 2\lambda\upsilon_z S_z$$
$$-\frac{V}{c}\left[\frac{\partial\rho}{\partial t} + \frac{\partial}{\partial x}(\rho\upsilon_x) + \frac{\partial}{\partial y}\left(\rho\upsilon_y\right) + \frac{\partial}{\partial z}(\rho\upsilon_z)\right]$$
$$-\frac{\rho}{c}\left(\frac{\partial V}{\partial t} + \upsilon_x\frac{\partial V}{\partial x} + \upsilon_y\frac{\partial V}{\partial y} + \upsilon_z\frac{\partial V}{\partial z}\right)$$
$$\frac{\partial\Phi^{0j}}{\partial x^j} = 2\lambda\left(-W + \upsilon_x S_x + \upsilon_y S_y + \upsilon_z S_z\right)$$
$$-\frac{V}{c}\left[\frac{\partial\rho}{\partial t} + \nabla\cdot(\rho\boldsymbol{\upsilon})\right] - \frac{\rho}{c}\left(\frac{\partial V}{\partial t} + \boldsymbol{\upsilon}\cdot\nabla V\right)$$

and from the equation of continuity, as well as equation (254), we get

$$\frac{\partial\Phi^{0j}}{\partial x^j} = 2\lambda\left(-W + \upsilon_x S_x + \upsilon_y S_y + \upsilon_z S_z\right) + \frac{\rho}{c}(\boldsymbol{\upsilon}\cdot\boldsymbol{\varepsilon})$$

and with equations (251) and (252) we get

$$\frac{\partial\Phi^{0j}}{\partial x^j} = 2\lambda\varepsilon_0\left[-\frac{1}{2}\boldsymbol{\varepsilon}^2 - \frac{1}{2}\boldsymbol{B}^2 + \frac{\upsilon_x^2}{c^2}\boldsymbol{\varepsilon}^2 - \upsilon_x\varepsilon_x\left(\frac{\boldsymbol{\upsilon}\cdot\boldsymbol{\varepsilon}}{c^2}\right) + \frac{\upsilon_y^2}{c^2}\boldsymbol{\varepsilon}^2\right.$$
$$\left.-\upsilon_y\varepsilon_y\left(\frac{\boldsymbol{\upsilon}\cdot\boldsymbol{\varepsilon}}{c^2}\right) + \frac{\upsilon_z^2}{c^2}\boldsymbol{\varepsilon}^2 - \upsilon_z\varepsilon_z\left(\frac{\boldsymbol{\upsilon}\cdot\boldsymbol{\varepsilon}}{c^2}\right)\right] + \rho\left(\frac{\boldsymbol{\upsilon}\cdot\boldsymbol{\varepsilon}}{c}\right)$$

$$\frac{\partial\Phi^{0j}}{\partial x^j} = -2\lambda\varepsilon_0\left[-\frac{1}{2}\boldsymbol{\varepsilon}^2 - \frac{1}{2}c^2\boldsymbol{B}^2 + \frac{\upsilon_x^2+\upsilon_y^2+\upsilon_z^2}{c^2}\boldsymbol{\varepsilon}^2\right] -$$
$$-2\lambda\varepsilon_0\left[-(\boldsymbol{\upsilon}\cdot\boldsymbol{\varepsilon})\left(\upsilon_x\varepsilon_x + \upsilon_y\varepsilon_y + \upsilon_z\varepsilon_z\right)\right] + \rho\left(\frac{\boldsymbol{\upsilon}\cdot\boldsymbol{\varepsilon}}{c}\right)$$

and since it is $\upsilon_x^2 + \upsilon_y^2 + \upsilon_z^2 = c^2$ and also

$$\upsilon_x\varepsilon_x + \upsilon_y\varepsilon_y + \upsilon_z\varepsilon_z = \boldsymbol{\upsilon}\cdot\boldsymbol{\varepsilon},$$

we see that

$$\frac{\partial\Phi^{0j}}{\partial x^j} = -2\lambda\varepsilon_0\left[-\frac{1}{2}\boldsymbol{\varepsilon}^2 - \frac{1}{2}c^2\boldsymbol{B}^2 - \left(\frac{\boldsymbol{\upsilon}\cdot\boldsymbol{\varepsilon}}{c}\right)^2\right] + \rho\left(\frac{\boldsymbol{\upsilon}\cdot\boldsymbol{\varepsilon}}{c}\right).$$

From equation (235) we obtain

$$\frac{\partial\Phi^{0j}}{\partial x^j} = -2\lambda\varepsilon_0\frac{1}{2}\left(\frac{\boldsymbol{\upsilon}\cdot\boldsymbol{\varepsilon}}{c}\right)^2 + \rho\left(\frac{\boldsymbol{\upsilon}\cdot\boldsymbol{\varepsilon}}{c}\right)$$
$$\frac{\partial\Phi^{0j}}{\partial x^j} = \left(\frac{\boldsymbol{\upsilon}\cdot\boldsymbol{\varepsilon}}{c}\right)\left[\rho - \lambda\varepsilon_0\frac{\boldsymbol{\upsilon}\cdot\boldsymbol{\varepsilon}}{c}\right] = 0$$

since

$$\rho - \lambda\varepsilon_0\frac{\boldsymbol{\upsilon}\cdot\boldsymbol{\varepsilon}}{c} = 0. \tag{258}$$

Indeed, substituting the factor λ from equations (257), we have

$$\lambda\varepsilon_0\frac{\boldsymbol{\upsilon}\cdot\boldsymbol{\varepsilon}}{c} = -\frac{1}{1-\dfrac{\boldsymbol{\upsilon}\cdot\boldsymbol{u}}{c^2}}\frac{\partial q}{q\partial w}\varepsilon_0\frac{\boldsymbol{\upsilon}\cdot\boldsymbol{\varepsilon}}{c}$$
$$\lambda\varepsilon_0\frac{\boldsymbol{\upsilon}\cdot\boldsymbol{\varepsilon}}{c} = -\frac{1}{1-\dfrac{\boldsymbol{\upsilon}\cdot\boldsymbol{u}}{c^2}}\frac{\partial q}{q\partial w}\varepsilon_0\frac{\boldsymbol{\upsilon}}{c}(\boldsymbol{\varepsilon}_u + \boldsymbol{\varepsilon}_\alpha).$$

From equation (234) we have

$$\boldsymbol{\upsilon}\cdot\boldsymbol{\varepsilon}_\alpha = 0.$$

Hence, we see that

$$\lambda\varepsilon_0\frac{\boldsymbol{\upsilon}\cdot\boldsymbol{\varepsilon}}{c} = -\frac{\varepsilon_0}{1-\dfrac{\boldsymbol{\upsilon}\cdot\boldsymbol{u}}{c^2}}\frac{\partial q}{q\partial w}\frac{\boldsymbol{\upsilon}}{c}\boldsymbol{\varepsilon}_u$$

and with equation (208) for the electric field $\boldsymbol{\varepsilon}_u$ we get

$$\lambda\varepsilon_0\frac{\boldsymbol{\upsilon}\cdot\boldsymbol{\varepsilon}}{c} = -\frac{\varepsilon_0}{1-\dfrac{\boldsymbol{\upsilon}\cdot\boldsymbol{u}}{c^2}}\frac{\partial q}{q\partial w}\frac{q\left(1-\dfrac{u^2}{c^2}\right)}{4\pi\varepsilon_0 r^2\left(1-\dfrac{\boldsymbol{\upsilon}\cdot\boldsymbol{u}}{c^2}\right)^3}\frac{\boldsymbol{\upsilon}}{c}\left(\frac{\boldsymbol{\upsilon}}{c} - \frac{\boldsymbol{u}}{c}\right)$$

$$\lambda\varepsilon_0\frac{\boldsymbol{\upsilon}\cdot\boldsymbol{\varepsilon}}{c} = -\frac{\partial q}{\partial w}\frac{1-\dfrac{u^2}{c^2}}{4\pi r^2\left(1-\dfrac{\boldsymbol{\upsilon}\cdot\boldsymbol{u}}{c^2}\right)^4}\left(1-\frac{\boldsymbol{\upsilon}\cdot\boldsymbol{u}}{c}\right)$$

$$\lambda\varepsilon_0\frac{\boldsymbol{\upsilon}\cdot\boldsymbol{\varepsilon}}{c} = -\frac{\partial q}{\partial w}\frac{1-\dfrac{u^2}{c^2}}{4\pi r^2\left(1-\dfrac{\boldsymbol{\upsilon}\cdot\boldsymbol{u}}{c^2}\right)^3}.$$

Applying equation (205) we get

$$\lambda\varepsilon_0\frac{\boldsymbol{\upsilon}\cdot\boldsymbol{\varepsilon}}{c} = \rho$$
$$\rho - \lambda\varepsilon_0\frac{\boldsymbol{\upsilon}\cdot\boldsymbol{\varepsilon}}{c} = 0.$$

The validity of equation (256) for $i = 1, 2, 3$ is proven similarly.

In paragraph 4.5 we proved that the selfvariations are in agreement with the conservation of energy and momentum. The proof was done in two different ways: by direct calculation, and by applying the continuity equation. While it is of interest that the two different proofs, both lead to the conclusion that the selfvariations are compatible with the conservation principles of Physics, the calculation for the energy-momentum tensor was done for a completely different, and very substantial, reason. At macrocosmic scales, that is at large distances from the material particle, where equations

(151) and (152) hold, the energy-momentum tensor Φ^{ij}, as given by equation (254), indeed contains all the information about the energy content of spacetime. At microcosmic scales the equations of the theory of selfvariations highlight additional parameters about the energy content of spacetime. These parameters bring the quantum phenomena to the forefront.

4.8 The energy-momentum tensor of the generalized photon at macrocosmic scales

In this paragraph we shall study the energy-momentum tensor for the generalized photon that balances the selfvariation of the rest mass of the material particle. Using our notation the energy-momentum tensor is given by

$$\Phi^{ij} = \frac{D}{c^2}\begin{bmatrix} c^2 & cv_x & cv_y & cv_z \\ v_x c & v_x^2 & v_x v_y & v_x v_z \\ v_y c & v_y v_x & v_y^2 & v_y v_z \\ v_z c & v_z v_x & v_z v_y & v_z^2 \end{bmatrix} \tag{259}$$

with the energy density D given by equation (167).

We shall prove the conservation of energy and momentum as given by equation

$$\Phi^{ij}_{,j} = \frac{\partial \Phi^{ij}}{\partial x^j} = 0. \tag{260}$$

For $i = 0$ we have

$$\frac{\partial \Phi^{0j}}{\partial x^j} = \frac{\partial \Phi^{00}}{\partial x^0} + \frac{\partial \Phi^{01}}{\partial x^1} + \frac{\partial \Phi^{02}}{\partial x^2} + \frac{\partial \Phi^{03}}{\partial x^3}$$

$$\frac{\partial \Phi^{0j}}{\partial x^j} = \frac{\partial}{c\partial t}\left(\frac{Dc^2}{c^2}\right) + \frac{\partial}{\partial x}\left(\frac{D}{c^2}cv_x\right) + \frac{\partial}{\partial y}\left(\frac{D}{c^2}cv_y\right) + \frac{\partial}{\partial z}\left(\frac{D}{c^2}cv_z\right)$$

$$\frac{\partial \Phi^{0j}}{\partial x^j} = \frac{1}{c}\left[\frac{\partial D}{\partial t} + \nabla\cdot(D\boldsymbol{v})\right]$$

and with equation (168) we get

$$\frac{\partial \Phi^{0j}}{\partial x^j} = 0.$$

For $i = 1$ we have

$$\frac{\partial \Phi^{ij}}{\partial x^j} = \frac{\partial \Phi^{10}}{\partial x^0} + \frac{\partial \Phi^{11}}{\partial x^1} + \frac{\partial \Phi^{12}}{\partial x^2} + \frac{\partial \Phi^{13}}{\partial x^3}$$

$$\frac{\partial \Phi^{ij}}{\partial x^j} = \frac{1}{c^2}\left[\frac{\partial}{\partial t}(Dv_x) + \frac{\partial}{\partial x}(Dv_x v_x) + \frac{\partial}{\partial y}\left(Dv_x v_y\right)\right] + \frac{1}{c^2}\left[\frac{\partial}{\partial z}(Dv_x v_z)\right]$$

$$\frac{\partial \Phi^{ij}}{\partial x^j} = \frac{1}{c^2}\left[v_x\left(\frac{\partial D}{\partial t} + \nabla\cdot(D\boldsymbol{v})\right) + D\left(\frac{\partial v_x}{\partial t} + \boldsymbol{v}\cdot\nabla v_x\right)\right]$$

and with equation (168) we get

$$\frac{\partial \Phi^{ij}}{\partial x^j} = v_x\left(\frac{\partial D}{\partial t} + \nabla\cdot\boldsymbol{j}\right) + \frac{1}{c^2}D\left(\frac{\partial v_x}{\partial t} + \boldsymbol{v}\cdot\nabla v_x\right)$$

and with (185) we arrive at

$$\frac{\partial \Phi^{ij}}{\partial x^j} = \frac{D}{c^2}\left(\frac{\partial v_x}{\partial t} + \boldsymbol{v}\cdot\nabla v_x\right). \tag{261}$$

From equation (33) we have

$$\frac{\partial v_x}{\partial t} + \boldsymbol{v}\cdot\nabla v_x = \frac{\partial}{\partial t}(c\cos\delta) + \boldsymbol{v}\cdot\nabla(c\cos\delta)$$

$$\frac{\partial v_x}{\partial t} + \boldsymbol{v}\cdot\nabla v_x = -c\sin\delta\left(\frac{\partial \delta}{\partial t} + \boldsymbol{v}\cdot\nabla\delta\right)$$

and with equation (81)(b) we get

$$\frac{\partial v_x}{\partial t} + \boldsymbol{v}\cdot\nabla v_x = 0. \tag{262}$$

Combining equations (261) and (262), we see that

$$\frac{\partial \Phi^{ij}}{\partial x^j} = 0.$$

We can similarly prove the validity of equation (259) for $i = 2, 3$.

By comparing the results of the last two paragraphs we find substantial differences between the generalized photon that counterbalances the selfvariation of the electric charge and the generalized photon that counterbalances the selfvariation of the rest mass of the material particle. Within the energy-momentum tensor of the first, there appears the electromagnetic field, as expressed by the first matrix of the second part of equation (254). On the contrary, in the expression of the energy-momentum tensor of equation (259), no corresponding matrix appears. Therefore, the generalized photon counterbalancing the rest mass does not correspond to a kind of field with the structure and content of the electromagnetic field. Furthermore, by comparing the second matrix of equation (254) with the matrix of equation (259), we observe that in place of the potential V in the first, the factor c^2 appears in the second. These observations hold even if we formulate the equations for a non-inertial reference frame (we have already suggested a way for formulating the equations in non-inertial reference frames). By careful observation of the equations appearing in paragraphs 4.2, 4.3 and 4.4, we realize that the difference in the "behavior" of the couples $(\rho, \boldsymbol{j})$ and $(D, \boldsymbol{J})$ is the result of the different way the electric charge and the energy transform according to Lorentz-Einstein. It is exactly this difference that is captured on tensors (254) and (259). The generalized photon gives us the exact mechanism of transport of energy and momentum from one material particle to the other. At the same time, it highlights the similarities and differences between the electromagnetic and the gravitational interaction.

We could call the generalized photon that counterbalances the selfvariation of the rest mass by a different name. In any case it is obvious when we refer to the electric charge and when we refer to the rest mass. We shall, therefore, keep the name "generalized photon" for both cases.

The observation we made at the end of the previous paragraph regarding the tensor given by equation (254), also holds for tensor (259). It is valid at macrocosmic scales. At microcosmic scales, further parameters emerge from the theory of selfvariations, which cannot be given by the energy-momentum tensor.

4.9 The internality of the universe to the measurement procedure

The selfvariations hypothesis brings to the foreground the "internality of the Universe to the measurement procedure". Usually, in order to measure a physical quantity, we define as unit an arbitrary quantity with which we compare other physical quantities of the same kind. If the defined unit of measurement depends on the rest mass or the electric charge, then it is itself subject to the selfvariations. This fact must be taken into account every time we perform a measurement.

The photon does not have rest mass or electric charge and is, therefore, not affected by the selfvariations. The evidence we have suggests that the selfvariations take place at extremely slow rates. Therefore, the first consequence of the selfvariations we expect to observe is the following: photons with great lifetimes will be measured to have less energy than expected.

The extremely slow rate of evolution of the selfvariations, combined with the "internality of the Universe to the measurement procedure", do not allow their immediate observation in the laboratory. In the laboratory we only observe the consequences of the selfvariations. These consequences are the potential fields and the quantum phenomena.

5 The quantitative determination of the selfvariations

5.1 Introduction

In the present paragraph we develop the main axis of the structure of the theory of selfvariations. We determine quantitatively the rate of evolution of the selfvariations, and formulate the *law of selfvariations*.

The law of selfvariations dominates from the microcosmic scales up to the observations we conduct billions of light years away. It reveals the causes of quantum phenomena, while it contains as physical information the totality of the cosmological observational data. At the same time, it sets the path for understanding the interactions between material particles.

The equations resulting from the law of selfvariations are of fundamental nature for the science of Physics and the related Physical Sciences. They contain a large amount of physical information, which permits the full understanding of physical reality.

5.2 The law of selfvariations

The conclusions derived in the previous paragraphs refer to the surrounding spacetime of the material particle. These conclusions are grounded on the second proposition-axiom of the theory of selfvariations, which states that

$$dS^2 = 0. \tag{263}$$

This proposition is equivalent to the relation $\|\boldsymbol{v}\| = c$ which holds in every inertial system of reference.

In figure 4 the rest mass m_0 and the electric charge q of the material particle act at point $A(x, y, z, t)$ with the value they acquired at the moment $w = t - \frac{r}{c}$. Thus, we have that $m_0 = m_0(w)$ and $q = q(w)$. For the relevant calculations and proofs we have taken into consideration the axioms of the theory of selfvariations, but we have not yet defined the rate of evolution of their manifestation. In order to study the consequences of the selfvariations we have to determine quantitatively the first proposition-axiom of the theory of selfvariations.

Equation (263), combined with the first proposition-axiom of the selfvariations, leads directly to the concept of the *"generalized photon"*. The material particle emits generalized photons, and each generalized photon carries energy E and momentum $\boldsymbol{P}$, in order to counterbalance the change in energy and momentum that results from the selfvariations of the rest mass of the material particle. If the material particle also carries electric charge, then the generalized photon carries electric charge as well, in order to counterbalance the variation of the electric charge of the material particle due to the selfvariations.

The rate of evolution of the selfvariations is determined axiomatically with the help of the total energy E_s and the total momentum $\boldsymbol{P}_s$, which is emitted simultaneously and in all directions by the material particle, according to the following proposition-axiom: "The rest mass m_0 and the electric charge q of every material particle vary according to the action of the operators

$$\frac{\partial}{\partial t} \to -\frac{i}{\hbar} E_s \qquad \nabla \to \frac{i}{\hbar} \boldsymbol{P}_s \tag{264}$$

where E_s and $\boldsymbol{P}_s$ denote the total energy and total momentum of the generalized photons emitted simultaneously by the material particle in all directions, and $\hbar = \frac{h}{2\pi}$, where h is Planck's constant".

Stated in the form of equations, relations (264) can be written as

$$\begin{aligned} \frac{\partial m_0}{\partial t} &= -\frac{i}{\hbar} E_s m_0 \\ \nabla m_0 &= \frac{i}{\hbar} \boldsymbol{P}_s m_0 \end{aligned} \tag{265}$$

and

$$\frac{\partial q}{\partial t} = -\frac{i}{\hbar} E_s q$$
$$\nabla q = \frac{i}{\hbar} \boldsymbol{P}_s q. \tag{266}$$

In equations (265) and (266) we use the same symbol for the energy E_s and the momentum $\boldsymbol{P}_s$. But these are not the same physical quantities. In equations (265) the energy E_s and the momentum $\boldsymbol{P}_s$ counterbalance the consequences of the selfvariations of the rest mass. In equations (266) they counterbalance the consequences of the selfvariations of the electric charge. Later, we shall modify equation (266) in order to make this difference transparent.

The emission of generalized photons by the material particle comes about, initially, as a consequence of the principles of conservation of energy, momentum and electric charge. The operators given in relations (264) determine the relation between the material particle and the generalized photons, independently from the principles of conservation. Equations (265) and (266) express in a quantitative manner the *law of selfvariations.*

According to the law of selfvariations the rest mass m_0 and the electric charge q are functions of time t, as well as of the position of the material particle

$$m_0 = m_0(X_p, Y_p, Z_p, t)$$
$$q = q(X_p, Y_p, Z_p, t). \tag{267}$$

The dependence of the rest mass and the electric charge, not only on time, but also on the spatial position, is to be expected. Even if in some inertial frame of reference they only depend on time, in another inertial frame of reference they will also depend on the position, according to the Lorentz-Einstein transformations.

From equation (200), and for $\boldsymbol{u} = 0$, we take that $\boldsymbol{P}_s = 0$, so that the second equation of the couple of equations (265) gives $\nabla m_0 = \mathbf{0}$, whereas the first equation can be written as

$$\frac{dm_0}{dt} = -\frac{i}{\hbar} E_0 m_0$$
$$\dot{m} = -\frac{i}{\hbar} E_0 m_0$$
$$E_0 = i\hbar \frac{\dot{m}_0}{m_0}. \tag{268}$$

Here, we denote the differentiation with respect to time by $(\bullet)$, and we set $E_s = E_0$ (the necessity of denoting $E_s = E_0$ will become apparent later on).

Furthermore, from the principle of conservation of energy at the instant of emission of the generalized photons, we obtain that

$$(m_0 c^2 + E_0)^{\bullet} = 0. \tag{269}$$

Combining equations (268) and (269) we arrive at equation

$$\left(m_0 c^2 + i\hbar \frac{\dot{m}_0}{m_0} \right)^{\bullet} = 0. \tag{270}$$

Equation (270) both contains as physical information, and justifies, the whole corpus of the current cosmological observational data, as described in paragraph 7.

5.3 The "percentage function" Φ

The law of selfvariations expresses the total interaction of the generalized photons, which are emitted simultaneously by the material particle, with its rest mass and electric charge. However, in a particular direction $\frac{\boldsymbol{v}}{c}$, the material particle emits generalized photons of energy E and momentum $\boldsymbol{P}$. Therefore, we have to derive quantitatively the partial contribution of a single generalized photon of energy E and momentum $\boldsymbol{P}$ to the law of selfvariations.

We have to answer the following question:

"Which mathematical equation correlates the energy E and the momentum $\boldsymbol{P}$ of a single generalized photon emitted towards a particular direction $\frac{\boldsymbol{v}}{c}$, to the selfvariations of the rest mass m_0 and the electric charge q of the material particle?"

Thus, we are seeking the form of equations (265) and (266) that correspond to a single generalized photon.

Based on the law of selfvariations, the answer to this physical problem can only be given by the following statement:

"The partial contribution of a single generalized photon to the selfvariations of the rest mass m_0 and the electric charge q of the material particle is given by any mathematical expression which agrees with the operators defined in equations (264). If we sum the contributions of the single generalized photons towards all directions, during their simultaneous emission by the material particle, we have to end up with the equations given in (265) and (266)".

Considering this physical problem from its mathematical aspect, we can choose arbitrarily any mathematical expression giving the partial contribution of a single generalized photon according to the law of selfvariations, which satisfies the operators (264). Then, we can compare the results obtained by our particular choice with physical reality. On the other hand, we can choose the mathematical expression taking into account some specific *physical* criteria beforehand.

A fundamental case for the partial contribution of a generalized photon according to the law of selfvariations arises from the following observation: A single generalized photon counterbalances only a percentage of the total energy, momentum and electric charge that result from the selfvariations. Therefore, we must examine whether the contribution of a single generalized photon to the law of selfvariations is correlated with a *percentage* Φ of the rest mass m_0 and electric charge q. In this case, the partial contribution to the law

of selfvariations for a single generalized photon of energy E and momentum $\boldsymbol{P}$ will be given by the set of equations

$$\begin{aligned} \frac{\partial(\Phi m_0)}{\partial t} &= -\frac{i}{\hbar} E m_0 \\ \nabla(\Phi m_0) &= \frac{i}{\hbar} \boldsymbol{P} m_0 \end{aligned} \tag{271}$$

$$\begin{aligned} \frac{\partial(\Phi q)}{\partial t} &= -\frac{i}{\hbar} E q \\ \nabla(\Phi q) &= \frac{i}{\hbar} \boldsymbol{P} q. \end{aligned} \tag{272}$$

Summing in all directions of emission of generalized photons in the first equation of the set of equations (271), we obtain relations

$$\begin{aligned} \sum \frac{\partial(\Phi m_0)}{\partial t} &= -\frac{i}{\hbar} \sum E m_0 \\ \frac{\partial}{\partial t}\left(\sum \Phi m_0\right) &= -\frac{i}{\hbar} m_0 \sum E \\ \frac{\partial}{\partial t}\left(m_0 \sum \Phi\right) &= -\frac{i}{\hbar} m_0 \sum E. \end{aligned}$$

Since it holds that $\sum E = E_s$ and the total percentage of the contributions is 1, that is $\sum \Phi = 1$, we get

$$\frac{\partial m_0}{\partial t} = -\frac{i}{\hbar} m_0 E_s.$$

This is the first equation of the set of equations (265).

Also, from the second equation of the set of equations (271) we obtain relations

$$\begin{aligned} \sum \nabla(\Phi m_0) &= \frac{i}{\hbar} \sum \boldsymbol{P} m_0 \\ \nabla\left(\sum (\Phi m_0)\right) &= \frac{i}{\hbar} m_0 \sum \boldsymbol{P} \\ \nabla\left(m_0 \sum \Phi\right) &= \frac{i}{\hbar} m_0 \sum \boldsymbol{P}. \end{aligned}$$

Since $\sum \Phi = 1$ and $\sum \boldsymbol{P} = \boldsymbol{P}_s$, we see that

$$\nabla m_0 = \frac{i}{\hbar} m_0 \boldsymbol{P}_s.$$

This is the second of the equations given in (265).

We can perform the same procedure for equations (272) as well. Therefore, a single generalized photon can contribute to the selfvariation with a percentage Φ of the rest mass or the electric charge, and then this contribution is expressed by equations (271) and (272).

From equations (271) we obtain

$$\begin{aligned} \Phi \frac{\partial m_0}{\partial t} + m_0 \frac{\partial \Phi}{\partial t} &= -\frac{i}{\hbar} E m_0 \\ \Phi \nabla m_0 + m_0 \nabla \Phi &= \frac{i}{\hbar} \boldsymbol{P} m_0. \end{aligned}$$

From equations (169) we also obtain

$$\begin{aligned} \Phi \frac{1}{1 - \dfrac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}} \frac{\partial m_0}{\partial w} + m_0 \frac{\partial \Phi}{\partial t} &= -\frac{i}{\hbar} E m_0 \\ -\Phi \frac{1}{1 - \dfrac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}} \frac{\partial m_0}{c\partial w} \frac{\boldsymbol{v}}{c} + m_0 \nabla \Phi &= \frac{i}{\hbar} \boldsymbol{P} m_0. \end{aligned}$$

Eliminating from the equations the quantity m_0, we obtain

$$\begin{aligned} \frac{1}{1 - \dfrac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}} \frac{\partial m_0}{m_0 \partial w} + \frac{\partial \Phi}{\partial t} &= -\frac{i}{\hbar} E \\ -\frac{1}{1 - \dfrac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}} \frac{\partial m_0}{m_0 c \partial w} \frac{\boldsymbol{v}}{c} + \nabla \Phi &= \frac{i}{\hbar} \boldsymbol{P}. \end{aligned}$$

Finally, we arrive at the set of equations

$$\begin{aligned} E &= \Phi \frac{i\hbar}{1 - \dfrac{\boldsymbol{v}\boldsymbol{u}}{c^2}} \frac{\partial m_0}{m_0 \partial w} + i\hbar \frac{\partial \Phi}{\partial t} \\ \boldsymbol{P} &= \Phi \frac{i\hbar}{1 - \dfrac{\boldsymbol{v}\boldsymbol{u}}{c^2}} \frac{\partial m_0}{m_0 c \partial w} \frac{\boldsymbol{v}}{c} - i\hbar \nabla \Phi. \end{aligned} \tag{273}$$

The function Φ can be any mathematical function, defined on the material particle and obeying relation

$$\sum \Phi = 1 \tag{274}$$

However, it has to be considered a function depending on the direction in space, since this is implied by the summation given in equation (274).

According to the operators defined in (264), the continuous evolution of the selfvariations with the passage of time is assured by the condition

$$E_s \neq 0. \tag{275}$$

This condition is a straightforward consequence of the first proposition-axiom of the theory of selfvariations.

We are seeking now to derive the relation between the total momentum $\boldsymbol{P}_s$ and the total energy E_s. According to equation (200) this relation can be written as

$$\boldsymbol{P}_s = E_s \frac{\boldsymbol{u}}{c^2}. \tag{276}$$

Here, $\boldsymbol{u}$ denotes the velocity of the material particle at the moment of the emission of the generalized photons.

This relation has to be reconsidered for the following reason: During the proof of this relation in paragraph 4.3, we have taken into consideration equation (168), that is equation

$$\boldsymbol{J} = D \frac{\boldsymbol{v}}{c^2}.$$

This equation presupposes the validity of the condition

$$\boldsymbol{P} = E\frac{\boldsymbol{v}}{c^2} \tag{277}$$

for every single generalized photon emitted towards any direction defined by $\frac{\boldsymbol{v}}{c}$, as depicted in figure 6. However, equations (273) reveal a more complex, and certainly different relation, between the momentum $\boldsymbol{P}$ and the energy E of a single generalized photon. Therefore, we have to reconsider the validity of equation (276), since we cannot base its proof on equation (277). As we shall see immediately, equation (276) is of general validity, and is compatible with the set of equations (273).

We consider a material point particle at rest, as depicted in figure 7. In order for this particle to remain at rest, the total momentum emitted simultaneously and towards all directions has to vanish, that is

$$\boldsymbol{P}_s^{'} = \boldsymbol{0}. \tag{278}$$

If the case were different, the material particle would undergo an arbitrary motion, as a consequence of the principle of conservation of momentum. From equation (278), and from the set of transformations (164) for the total energy E_s and the total momentum $\boldsymbol{P}_s$, we arrive at equation (276). Thus, we have

$$\begin{aligned} E_s &= \gamma\left(\mathrm{E}_s^{'} + uP_{sx}^{'}\right) \\ P_{sx} &= \gamma\left(P_{sx}^{'} + \frac{u}{c^2}\mathrm{E}_s^{'}\right) \\ P_{sy} &= P_{sy}^{'} \\ P_{sz} &= P_{sz}^{'}. \end{aligned}$$

Since, according to equation (278) it holds that

$$\left(P_{sx}^{'}, P_{sy}^{'}, P_{sz}^{'}\right) = (0,0,0)\,,$$

we obtain the following relations

$$\begin{aligned} E_s &= \gamma\mathrm{E}_s^{'} \\ P_{sx} &= \gamma\mathrm{E}_s^{'}\frac{u}{c^2} \\ P_{sy} &= 0 \\ P_{sz} &= 0. \end{aligned}$$

We also have that $\boldsymbol{u} = \begin{bmatrix} u \\ 0 \\ 0 \end{bmatrix}$, thus we obtain

$$\begin{aligned} E_s &= \gamma\mathrm{E}_s^{'} \\ \boldsymbol{P}_s &= \gamma\mathrm{E}_s^{'}\frac{\boldsymbol{u}}{c^2}. \end{aligned}$$

Finally, we have

$$\begin{aligned} E_s &= \gamma\mathrm{E}_s^{'} \\ \boldsymbol{P}_s &= E_s\frac{\boldsymbol{u}}{c^2}. \end{aligned}$$

This is equation (276). Furthermore, we also obtain equation

$$E_s = \gamma\mathrm{E}_s^{'} = \gamma E_0 = \frac{E_0}{\sqrt{1-\dfrac{u^2}{c^2}}}. \tag{279}$$

Here, we denote

$$\mathrm{E}_s^{'} = E_0. \tag{280}$$

A material particle at rest can emit generalized photons of different energies for different directions. If the generalized photons emitted in opposite directions have opposite momenta, the material particle will remain at rest. But the momentum of a generalized photon can also be balanced by two other generalized photons emitted towards appropriate directions and with appropriate energies. In reality, there is an infinite number of combinations of emmision of generalized photons, with infinite combinations of energies and directions of emission. In each of these cases where equation (278) holds, the particle remains at rest. The case of emission of identical generalized photons in all directions by a material particle at rest is only one among the infinite number of cases satisfying equation (278).

Therefore, by rotating the unit vector $\frac{\boldsymbol{v}^{'}}{c}$ around the point particle at rest, as depicted in figure 7, we expect a change in the energy of the generalized photons. Exactly this is shown by equations (273), while at the same time they highlight the factors defining the energy and momentum of each single generalized photon.

5.4 The accompanying particle

In the previous paragraph we proved equations (276) and (279):

$$\begin{aligned} \boldsymbol{P}_s &= E_s\frac{\boldsymbol{u}}{c^2} \\ E_s &= \frac{E_0}{\sqrt{1-\dfrac{u^2}{c^2}}}. \end{aligned} \tag{281}$$

Equations (281) show that the total energy and momentum emitted simultaneously and in all directions by the material particle behaves as a particle moving with velocity $\boldsymbol{u}$, and accompanying the material particle. There is a definite correspondence between equations (281) and equations

$$\begin{aligned} \boldsymbol{P} &= m\boldsymbol{u} \\ m &= \frac{m_0}{\sqrt{1-\dfrac{u^2}{c^2}}} \end{aligned}$$

which give the momentum $\boldsymbol{P}$ and the mass m of the material particle.

According to equations (281), the accompanying particle has rest energyE_0. This is the rest energy E_s' from equation (280). Therefore, the accompanying particle has a rest mass given by $\frac{E_0}{c^2}$.

According to the first proposition-axiom of the theory of selfvariations, the rest mass $\frac{E_0}{c^2}$ of the accompanying particle changes with the passage of time. Hence, we seek the counterparts of equations (265), which define the rate of change of the rest mass $\frac{E_0}{c^2}$, or equivalently the rest energy E_0. As such, we obtain the corresponding form of equations (265)

$$\begin{aligned}\frac{\partial E_0}{\partial t} &= \frac{i}{\hbar}mc^2E_0 = \frac{i}{\hbar}\gamma m_0c^2E_0 = \frac{i}{\hbar}\frac{m_0c^2}{\sqrt{1-\frac{u^2}{c^2}}}E_0\\ \nabla E_0 &= -\frac{i}{\hbar}m\boldsymbol{u}E_0 = -\frac{i}{\hbar}\gamma m_0\boldsymbol{u}E_0 = -\frac{i}{\hbar}\frac{m_0}{\sqrt{1-\frac{u^2}{c^2}}}\boldsymbol{u}E_0.\end{aligned} \tag{282}$$

Equations (265) describe the effect of the generalized photons on the rest mass of the material particle. In nature, though, effects are always mutual. Hence, just as the generalized photons affect the material particle, the material particle in turn affects the generalized photons, and these mutual interactions must occur in the framework of the same physical law. Therefore, from the outset the issue arises of the existence of a rest mass concealed within the operators (264), and of a corresponding equation symmetrical to (265). The quest for the partial contribution of a single generalized photon to the law of selfvariations revealed the existence of the rest mass $\frac{E_0}{c^2}$ and equations (282). The existence of the rest mass $\frac{E_0}{c^2}$ is predicted by the initial equations we formulated for the macrocosmic scales, through equation (200).

A large part of the predictions of the theory of selfvariations can be made without the aid of equations (282). For example, the justification of the observational cosmological data can be obtained from (270), which is proven independently without resorting to equations (282). The same holds for equations (273). However, the accompanying particle is a direct consequence of the selfvariations. Indeed, if we combine the second of equations (281) with relation (275) we see immediately that

$$E_0 \neq 0. \tag{283}$$

The rest mass $\frac{E_0}{c^2}$ of the accompanying particle cannot vanish. Therefore, in order to study the consequences of the selfvariations in their totality, we have to take into account the existence and the properties of the accompanying particle. In nature there is always the system "material particle-accompanying particle".

Let M_0 be the rest mass of the system "material particle-accompanying particle", given by

$$M_0 = m_0 + \frac{E_0}{c^2}. \tag{284}$$

We have that

$$\begin{aligned}\frac{\partial M_0}{\partial t} &= \frac{\partial}{\partial t}\left(m_0 + \frac{E_0}{c^2}\right)\\ \frac{\partial M_0}{\partial t} &= \frac{\partial m_0}{\partial t} + \frac{\partial E_0}{c^2\partial t}.\end{aligned}$$

Using the first equations of the sets of equations given in (265) and (266), we obtain relation

$$\frac{\partial M_0}{\partial t} = -\frac{i}{\hbar}E_sm_0 + \frac{i}{\hbar}\gamma m_0E_0.$$

And using equation (279) we get

$$\begin{aligned}\frac{\partial M_0}{\partial t} &= -\frac{i}{\hbar}\gamma E_0m_0 + \frac{i}{\hbar}\gamma m_0E_0\\ \frac{\partial M_0}{\partial t} &= 0.\end{aligned} \tag{285}$$

Similarly, using the second equations of the sets of equations (265) and (282) we have that

$$\nabla M_0 = \boldsymbol{0}. \tag{286}$$

From equations (285) and (286) we conclude that the rest mass M_0 of the system "material particle-accompanying particle" is a physical quantity not affected by the process of the selfvariations. Therefore, we can use the rest mass M_0 and the rest energy M_0c^2 as a unit of measurement of mass and energy, respectively. By this approach we circumvent the methodological problems stemming from the principle of the *"internality of the universe with respect to the measurement procedure"*, as stated in paragraph 4.9.

The quantitative mathematical description of physical reality depends on our ability to include in our equations the consequences of the internality of the universe to the measurement procedure. In the macrocosmic scales there is a very simple way to accomplish this, as described in paragraph 7. In the microcosmic scale we use as units of measurement of mass and energy the quantities M_0 and M_0c^2, respectively.

We rewrite now equations (265) in the form

$$\begin{aligned}\frac{\partial}{\partial t}\left(\frac{m_0}{M_0}\right) &= -\frac{i}{\hbar}E_s\left(\frac{m_0}{M_0}\right)\\ \nabla\left(\frac{m_0}{M_0}\right) &= \frac{i}{\hbar}\boldsymbol{P}_s\left(\frac{m_0}{M_0}\right).\end{aligned} \tag{287}$$

These equations have the exact same physical content as equations (265). They give the rate of change of the rest mass m_0, since the rest mass M_0 is not affected by the selfvariations, according to equations (285) and (286). At the same

time, these equations highlight the action of the operators (264) on the complex number $\frac{m_0}{M_0} \in \mathbb{C}$, since the complex unit i appears within the expressions of the operators. The same procedure can be repeated for the case of equations (282) as well, by introducing the number $\frac{E_0}{M_0 c^2} \in \mathbb{C}$, and for the whole list of equations we have stated.

The accompanying particle has rest mass of magnitude $\frac{E_0}{c^2}$, which comes from the sum of the contributions of the generalized photons emitted simultaneously by the material particle. This is the physical content of equations (281). Therefore, the mechanism through which the selfvariations occur plays a fundamental role for the determination of the physical properties of the accompanying particle, and eventually for the physical properties of the actual system "material particle-accompanying particle".

5.5 The symmetrical law for the electric charge

From the study already conducted in paragraph 4.2 it follows that the generalized photons counterbalancing the selfvariation of the electric charge q carry electric charge. Therefore, the physical object resulting from their aggregation carries electric charge q_i.

The law of the selfvariations for the electric charge q is given by equations (266)

$$\begin{aligned} \frac{\partial q}{\partial t} &= -\frac{i}{\hbar} E_s q \\ \nabla q &= \frac{i}{\hbar} \boldsymbol{P}_s q. \end{aligned} \tag{288}$$

In these equations we denote with E_s and $\boldsymbol{P}_s$ the total energy and momentum emitted by the material particle simultaneously in all directions, and which counterbalances the variations in energy resulting from the selfvariation of the electric charge. Although we have kept the same notation, these quantites are not the same as the ones appearing in equations (265).

In order to repeat the study conducted for the rest mass for the case of the electric charge, we have to define the equations symmetrical to (266). That is, we have to formulate the counterparts of equations (266) for the electric charge q_i.

The law of selfvariations for the electric charge (288) has to be modified so that it will define the interaction of the electric charges q and q_i, exactly as the law stated in equation (265) determines the interaction of the rest masses m_0 and $\frac{E_0}{c^2}$. Therefore, the second part of equation (288) has to be expressed such that the electric charge q_i appears. This can only be accomplished by the introduction of an electric potential V through equation

$$E_s = V q_i. \tag{289}$$

With this notation, and taking into account equation (276), equations (288) can be written as

$$\begin{aligned} \frac{\partial q}{\partial t} &= -\frac{i}{\hbar} V q_i q \\ \nabla q &= \frac{i}{\hbar} V q_i \frac{\boldsymbol{u}}{c^2} q. \end{aligned} \tag{290}$$

Equations (290) and (288) have the same physical content, if and only if the electric potential V is independent of the selfvariations.

Starting from equation (290), we can also deduce all equations inferred in the previous paragraphs for the rest mass, except now for the electric charge. The proof follows similar paths, and we shall note repeat it here in full.

Firstly, it can be deduced that the potential V can be written in the form

$$V = \gamma V_0 = \frac{V_0}{\sqrt{1 - \dfrac{u^2}{c^2}}}. \tag{291}$$

The potential V_0 stays invariant under the action of the Lorentz-Einstein transformations, and is independent of the selfvariations. The corresponding expressions of equations (268) and (270) are

$$\begin{aligned} q_i V_0 &= i\hbar \frac{\dot{q}}{q} \\ \left(q + \frac{i\hbar}{V_0} \frac{\dot{q}}{q} \right)^{\bullet} &= 0. \end{aligned} \tag{292}$$

The corresponding equations to the ones given in (273) for the generalized photon, can be formulated as

$$\begin{aligned} E &= \Phi \frac{i\hbar}{1 - \dfrac{\boldsymbol{v} \cdot \boldsymbol{u}}{c^2}} \frac{\partial q}{q \partial w} + i\hbar \frac{\partial \Phi}{\partial t} \\ \boldsymbol{P} &= \Phi \frac{i\hbar}{1 - \dfrac{\boldsymbol{v} \cdot \boldsymbol{u}}{c^2}} \frac{\partial q}{qc \partial w} \frac{\boldsymbol{v}}{c} - i\hbar \nabla \Phi. \end{aligned} \tag{293}$$

The corresponding equations to the equations (282), that is, the corresponding form of the law expressed in (290), are

$$\begin{aligned} \frac{\partial q_i}{\partial t} &= \frac{i}{\hbar} \gamma V_0 q q_i \\ \nabla q_i &= -\frac{i}{\hbar} \gamma V_0 q \frac{\boldsymbol{u}}{c^2} q_i. \end{aligned} \tag{294}$$

The corresponding relation to relation (283) is

$$q_i \neq 0. \tag{295}$$

The corresponding expression of equation (284), that is, the electric charge Q of the system "material particle-accompanying particle" is

$$Q = q + q_i. \tag{296}$$

The corresponding equations to equations (285) and (286) take the form

$$\begin{aligned} \frac{\partial Q}{\partial t} &= 0 \\ \nabla Q &= \mathbf{0}. \end{aligned} \tag{297}$$

The electric charge Q is not affected by the selfvariations.

5.6 Fundamental study of the generalized photon

In paragraph 4 we studied the consequences of the selfvariations in the surrounding spacetime of the material particle. In that study we considered the validity of equation (168)

$$\boldsymbol{J} = D\frac{\boldsymbol{v}}{c^2}$$

which presupposes the validity of equation

$$\boldsymbol{P} = E\frac{\boldsymbol{v}}{c^2} \tag{298}$$

for the generalized photon.

We know by now that the energy E and the momentum $\boldsymbol{P}$ of the generalized photon are not correlated through the simple relation (298). For the generalized photon that results from the selfvariation of the rest mass, equations (273) hold

$$\begin{aligned} E &= \Phi\frac{i\hbar}{1-\dfrac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}}\frac{\partial m_0}{m_0\partial w} + i\hbar\frac{\partial\Phi}{\partial t} \\ \boldsymbol{P} &= \Phi\frac{i\hbar}{1-\dfrac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}}\frac{\partial m_0}{m_0 c\partial w}\frac{\boldsymbol{v}}{c} - i\hbar\nabla\Phi. \end{aligned} \tag{299}$$

For the generalized photon that results from the selfvariation of the electric charge, equations (293) hold

$$\begin{aligned} E &= \Phi\frac{i\hbar}{1-\dfrac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}}\frac{\partial q}{q\partial w} + i\hbar\frac{\partial\Phi}{\partial t} \\ \boldsymbol{P} &= \Phi\frac{i\hbar}{1-\dfrac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}}\frac{\partial q}{qc\partial w}\frac{\boldsymbol{v}}{c} - i\hbar\nabla\Phi. \end{aligned} \tag{300}$$

Equations (299) and (300) lead to a completely different relation from (298), between the energy E and the momentum $\boldsymbol{P}$ of a generalized photon.

We will study the generalized photon, as given in equations (299). The study of equations (300) is exactly the same.

The percentage-function Φ depends on the direction $\frac{\boldsymbol{v}}{c}$ and can, therefore, be written as $\Phi = \Phi(\delta,\omega)$, and can also depend on the moment, $w = t - \frac{r}{c}$, of emission of the generalized photon, so that

$$\Phi = \Phi(w,\delta,\omega). \tag{301}$$

From the first of equations (299) we have

$$E = \Phi\frac{i\hbar}{1-\dfrac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}}\frac{\partial m_0}{m_0\partial w} + i\hbar\frac{\partial\Phi}{\partial t}$$

and with equation (301), we get

$$E = \Phi\frac{i\hbar}{1-\dfrac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}}\frac{\partial m_0}{m_0\partial w} + i\hbar\left(\frac{\partial\Phi}{\partial w}\frac{\partial w}{\partial t} + \frac{\partial\Phi}{\partial\delta}\frac{\partial\delta}{\partial t} + \frac{\partial\Phi}{\partial\omega}\frac{\partial\omega}{\partial t}\right)$$

and with equations (11), (41) and (42) we get

$$\begin{aligned} E &= \Phi\frac{i\hbar}{1-\dfrac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}}\frac{\partial m_0}{m_0\partial w} + \\ &+ \frac{i\hbar}{1-\dfrac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}}\left(\frac{\partial\Phi}{\partial w} - \frac{\boldsymbol{u}\cdot\boldsymbol{\beta}}{r}\frac{\partial\Phi}{\partial\delta} - \frac{\boldsymbol{u}\cdot\boldsymbol{\gamma}}{r\sin\delta}\frac{\partial\Phi}{\partial\omega}\right). \end{aligned} \tag{302}$$

From the second of equations (299), we have

$$\boldsymbol{P} = \Phi\frac{i\hbar}{1-\dfrac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}}\frac{\partial m_0}{m_0 c\partial w}\frac{\boldsymbol{v}}{c} - i\hbar\nabla\Phi$$

and with equation (301) we get

$$\boldsymbol{P} = \Phi\frac{i\hbar}{1-\dfrac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}}\frac{\partial m_0}{m_0 c\partial w}\frac{\boldsymbol{v}}{c} - i\hbar\left(\frac{\partial\Phi}{\partial w}\nabla w + \frac{\partial\Phi}{\partial\delta}\nabla\delta + \frac{\partial\Phi}{\partial\omega}\nabla\omega\right)$$

and with equations (12), (51) and (52), we get

$$\begin{aligned} \boldsymbol{P} &= \Phi\frac{i\hbar}{1-\dfrac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}}\frac{\partial m_0}{m_0 c\partial w}\frac{\boldsymbol{v}}{c} + \\ &+ \frac{i\hbar}{1-\dfrac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}}\frac{\partial\Phi}{\partial w}\frac{\boldsymbol{v}}{c^2} - \frac{i\hbar}{r}\frac{\partial\Phi}{\partial\delta}\left(\frac{\boldsymbol{u}\cdot\boldsymbol{\beta}}{1-\dfrac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}}\frac{\boldsymbol{v}}{c^2} + \boldsymbol{\beta}\right) - \\ &- \frac{i\hbar}{r\sin\delta}\frac{\partial\Phi}{\partial\omega}\left(\frac{\boldsymbol{u}\cdot\boldsymbol{\gamma}}{1-\dfrac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}}\frac{\boldsymbol{v}}{c^2} + \boldsymbol{\gamma}\right). \end{aligned} \tag{303}$$

We now denote

$$\begin{aligned} E_i &= \Phi\frac{i\hbar}{1-\dfrac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}}\frac{\partial m_0}{m_0\partial w} \\ P_i &= \Phi\frac{i\hbar}{1-\dfrac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}}\frac{\partial m_0}{m_0 c\partial w}\frac{\boldsymbol{v}}{c} \end{aligned} \tag{304}$$

$$E_\Phi = i\hbar\frac{\partial\Phi}{\partial t} = \frac{i\hbar}{1-\dfrac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}}\frac{\partial\Phi}{\partial w} - \frac{i\hbar\boldsymbol{u}\cdot\boldsymbol{\beta}}{r\left(1-\dfrac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}\right)}\frac{\partial\Phi}{\partial\delta} - \frac{i\hbar\boldsymbol{u}\cdot\boldsymbol{\gamma}}{r\left(1-\dfrac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}\right)\sin\delta}\frac{\partial\Phi}{\partial\omega}$$

$$\boldsymbol{P}_\Phi = -i\hbar\nabla\Phi = \frac{i\hbar}{1-\dfrac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}}\frac{\partial\Phi}{\partial w}\frac{\boldsymbol{v}}{c^2} - \frac{i\hbar}{r}\frac{\partial\Phi}{\partial\delta}\left(\frac{\boldsymbol{u}\cdot\boldsymbol{\beta}}{1-\dfrac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}}\frac{\boldsymbol{v}}{c^2}+\boldsymbol{\beta}\right) - \frac{i\hbar}{r\sin\delta}\frac{\partial\Phi}{\partial\omega}\left(\frac{\boldsymbol{u}\cdot\boldsymbol{\gamma}}{1-\dfrac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}}\frac{\boldsymbol{v}}{c^2}+\boldsymbol{\gamma}\right). \tag{305}$$

With this notation, equations (302) and (303) can be written as

$$\begin{aligned} E &= E_i + E_\Phi \\ \boldsymbol{P} &= \boldsymbol{P}_i + \boldsymbol{P}_\Phi. \end{aligned} \tag{306}$$

Combining equations (302) and (303), we obtain relation

$$\boldsymbol{P} = E\frac{\boldsymbol{v}}{c^2} - \frac{i\hbar}{r}\frac{\partial\Phi}{\partial\delta}\boldsymbol{\beta} - \frac{i\hbar}{r\sin\delta}\frac{\partial\Phi}{\partial\omega}\boldsymbol{\gamma} \tag{307}$$

relating the energy E and momentum P of the generalized photon.

The energy-momentum pair $(E_\Phi, \boldsymbol{P}_\Phi)$ can be decomposed into three partial pairs

$$\begin{aligned} E_w &= \frac{i\hbar}{1-\dfrac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}}\frac{\partial\Phi}{\partial w} \\ \boldsymbol{P}_w &= \frac{i\hbar}{1-\dfrac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}}\frac{\partial\Phi}{\partial w}\frac{\boldsymbol{v}}{c^2} \end{aligned} \tag{308}$$

$$\begin{aligned} E_\delta &= -\frac{i\hbar\boldsymbol{u}\cdot\boldsymbol{\beta}}{r\left(1-\dfrac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}\right)}\frac{\partial\Phi}{\partial\delta} \\ \boldsymbol{P}_\delta &= -\frac{i\hbar}{r}\frac{\partial\Phi}{\partial\delta}\left(\frac{\boldsymbol{u}\cdot\boldsymbol{\beta}}{1-\dfrac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}}\frac{\boldsymbol{v}}{c^2}+\boldsymbol{\beta}\right) \end{aligned} \tag{309}$$

$$\begin{aligned} E_\omega &= -\frac{i\hbar\boldsymbol{u}\cdot\boldsymbol{\gamma}}{r\left(1-\dfrac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}\right)\sin\delta}\frac{\partial\Phi}{\partial\omega} \\ \boldsymbol{P}_\omega &= -\frac{i\hbar}{r\sin\delta}\frac{\partial\Phi}{\partial\omega}\left(\frac{\boldsymbol{u}\cdot\boldsymbol{\gamma}}{1-\dfrac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}}\frac{\boldsymbol{v}}{c^2}+\boldsymbol{\gamma}\right) \end{aligned} \tag{310}$$

$$\begin{aligned} E_\Phi &= E_w + E_\delta + E_\omega \\ \boldsymbol{P}_\Phi &= \boldsymbol{P}_w + \boldsymbol{P}_\delta + \boldsymbol{P}_\omega. \end{aligned} \tag{311}$$

It is easy to prove that, in the case of constant-speed motion with velocity $\boldsymbol{u} = \begin{bmatrix} u \\ 0 \\ 0 \end{bmatrix}$, each of the energy-momentum pairs $(E_i, \boldsymbol{P}_i)$, $(E_w, \boldsymbol{P}_w)$, $(E_\delta, \boldsymbol{P}_\delta)$, $(E_\omega, \boldsymbol{P}_\omega)$ transforms autonomously, independently of the rest, according to the Lorentz-Einstein transformations. Furthermore, an invariant amount of energy corresponds to each pair.

We shall calculate the four invariant amounts of energy. In the same way, we can prove the independent Lorentz-Einstein transformations of the four energy-momentum pairs.

From equation (305) we have

$$E_i^2 - c^2\boldsymbol{P}_i^2 = E_i^2 - c^2E_i^2\left(\frac{\boldsymbol{v}}{c^2}\right)^2$$

and since $\boldsymbol{v}^2 = c^2$, we get

$$E_i^2 - c^2\boldsymbol{P}_i^2 = 0. \tag{312}$$

From equation (308) we have

$$E_w^2 - c^2\boldsymbol{P}_w^2 = E_w^2 - c^2E_w^2\left(\frac{\boldsymbol{v}}{c^2}\right)^2$$

and from $\boldsymbol{v}^2 = c^2$, we get

$$E_w^2 - c^2\boldsymbol{P}_w^2 = 0. \tag{313}$$

From equation (309) we have

$$E_\delta^2 - c^2\boldsymbol{P}_\delta^2 = -\frac{\hbar^2(\boldsymbol{u}\cdot\boldsymbol{\beta})^2}{r^2\left(1-\dfrac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}\right)^2}\left(\frac{\partial\Phi}{\partial\delta}\right)^2 + \frac{c^2\hbar^2}{r^2}\left(\frac{\partial\Phi}{\partial\delta}\right)^2\left[\frac{\boldsymbol{u}\cdot\boldsymbol{\beta}}{1-\dfrac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}}\frac{v}{c^2}+\boldsymbol{\beta}\right]^2$$

and since it is $\boldsymbol{u}\cdot\boldsymbol{\beta} = 0, \boldsymbol{v}^2 = c^2, \boldsymbol{\beta}^2 = 1$, we get

$$E_\delta^2 - c^2\boldsymbol{P}_\delta^2 = -\frac{\hbar^2(\boldsymbol{u}\cdot\boldsymbol{\beta})^2}{r^2\left(1-\dfrac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}\right)^2}\left(\frac{\partial\Phi}{\partial\delta}\right)^2 + \frac{\hbar^2(\boldsymbol{u}\cdot\boldsymbol{\beta})^2}{r^2\left(1-\dfrac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}\right)^2}\left(\frac{\partial\Phi}{\partial\delta}\right)^2 + \left(\frac{c\hbar}{r}\frac{\partial\Phi}{\partial\delta}\right)^2$$

$$E_\delta^2 - c^2\boldsymbol{P}_\delta^2 = \left(\frac{c\hbar}{r}\frac{\partial\Phi}{\partial\delta}\right)^2. \tag{314}$$

Similarly, from equations (310) we get

$$E_\omega^2 - c^2\boldsymbol{P}_\omega^2 = \left(\frac{c\hbar}{r\sin\delta}\frac{\partial\Phi}{\partial\omega}\right)^2. \tag{315}$$

From the transformations (127) we get

$$\frac{c\hbar}{r'}\frac{\partial\Phi}{\partial\delta'} = \frac{c\hbar}{\gamma r\left(1-\frac{u}{c}\cos\delta\right)}\frac{\partial\Phi}{\partial\delta}\gamma\left(1-\frac{u}{c}\cos\delta\right)$$

$$\frac{c\hbar}{r'}\frac{\partial\Phi}{\partial\delta'} = \frac{c\hbar}{\gamma r}\frac{\partial\Phi}{\partial\delta}. \tag{316}$$

Therefore, the second part of equation (314) remains invariant according to the Lorentz-Einstein transformations.

From transformations (124) and (127) we have

$$\frac{c\hbar}{r'\sin\delta'}\frac{\partial\Phi}{\partial\omega'} = \frac{c\hbar}{\gamma r\left(1-\frac{u}{c}\cos\delta\right)}\frac{\sin\delta}{\gamma\left(1-\frac{u}{c}\cos\delta\right)}\frac{\partial\Phi}{\partial\omega}$$

$$\frac{c\hbar}{r'\sin\delta'}\frac{\partial\Phi}{\partial\omega'} = \frac{c\hbar}{r\sin\delta}\frac{\partial\Phi}{\partial\omega}. \tag{317}$$

Therefore, the second part of equation (315) remains invariant under the Lorentz-Einstein transformations.

From equation (307) we can calculate the total invariant energy of the generalized photon

$$E^2 - c^2\boldsymbol{P}^2 = E^2 - c^2\left(E\frac{\boldsymbol{\upsilon}}{c^2} - \frac{i\hbar}{r}\frac{\partial\Phi}{\partial\delta}\boldsymbol{\beta} - \frac{i\hbar}{r\sin\delta}\frac{\partial\Phi}{\partial\omega}\boldsymbol{\gamma}\right)^2$$

and taking into consideration that the set of vectors $\frac{\upsilon}{c}, \boldsymbol{\beta}, \boldsymbol{\gamma}$ constitute an orthonormal basis, we get

$$E^2 - c^2\boldsymbol{P}^2 = E^2 - E^2 + \left(\frac{c\hbar}{r}\frac{\partial\Phi}{\partial\delta}\right)^2 + \left(\frac{c\hbar}{r\sin\delta}\frac{\partial\Phi}{\partial\omega}\right)^2$$

$$E^2 - c^2\boldsymbol{P}^2 = \left(\frac{c\hbar}{r}\frac{\partial\Phi}{\partial\delta}\right)^2 + \left(\frac{c\hbar}{r\sin\delta}\frac{\partial\Phi}{\partial\omega}\right)^2. \tag{318}$$

According to equations (316) and (317), the second part of equation (318) remains invariant under the Lorentz-Einstein transformations.

We will now prove that:

"In the case of constant-speed motion with velocity

$$\boldsymbol{u} = \begin{bmatrix} u \\ 0 \\ 0 \end{bmatrix},$$

pairs $(E_i, \boldsymbol{P}_i)$, $(E_w, \boldsymbol{P}_w)$ correspond to a flow of energy and momentum into the surrounding spacetime. On the contrary, pairs $(E_\delta, \boldsymbol{P}_\delta)$ and $(E_\omega, \boldsymbol{P}_\omega)$ correspond to a redistribution of energy and momentum in the surrounding spacetime".

From equation (109) together with the second of equations (308), we get

$$\boldsymbol{P}_i\cdot\boldsymbol{R} = \frac{i\hbar}{1-\frac{\boldsymbol{\upsilon}\cdot\boldsymbol{u}}{c^2}}\frac{\partial m_0}{m_0\partial w}\frac{\boldsymbol{\upsilon}}{c^2}r\left(\frac{\boldsymbol{\upsilon}}{c}-\frac{\boldsymbol{u}}{c}\right)$$

$$\boldsymbol{P}_i\cdot\boldsymbol{R} = \frac{i\hbar}{1-\frac{\boldsymbol{\upsilon}\cdot\boldsymbol{u}}{c^2}}\frac{\partial m_0}{m_0 c\partial w}r\left(1-\frac{\boldsymbol{\upsilon}\cdot\boldsymbol{u}}{c^2}\right)$$

$$\boldsymbol{P}_i\cdot\boldsymbol{R} = i\hbar r\frac{\partial m_0}{m_0 c\partial w}$$

$$\boldsymbol{P}_i\cdot\frac{\boldsymbol{R}}{r} = i\hbar\frac{\partial m_0}{m_0 c\partial w}. \tag{319}$$

Similarly, from equation (109) together with the second of equations (308) we get

$$\boldsymbol{P}_w\cdot\frac{\boldsymbol{R}}{r} = i\hbar\frac{\partial\Phi}{c\partial w}. \tag{320}$$

We conclude that both the momentum $\boldsymbol{P}_i$, as well as the momentum $\boldsymbol{P}_w$, have a component along the direction of vector $\boldsymbol{R}$, as depicted in Figure 6.

Combining equation (109) with the second of equations (309), we get

$$\boldsymbol{P}_\delta\cdot\boldsymbol{R} = -\frac{i\hbar}{r}\frac{\partial\Phi}{\partial\delta}\left(\frac{\boldsymbol{u}\cdot\boldsymbol{\beta}}{1-\frac{\boldsymbol{\upsilon}\cdot\boldsymbol{u}}{c^2}}\frac{\boldsymbol{\upsilon}}{c^2}+\boldsymbol{\beta}\right)r\left(\frac{\boldsymbol{\upsilon}}{c}-\frac{\boldsymbol{u}}{c}\right)$$

and since $\boldsymbol{\upsilon}^2 = c^2$ and $\boldsymbol{\upsilon}\cdot\boldsymbol{\beta} = 0$, we obtain

$$\boldsymbol{P}_\delta\cdot\boldsymbol{R} = -i\hbar\frac{\partial\Phi}{\partial\delta}\left(\frac{\boldsymbol{u}\cdot\boldsymbol{\beta}}{1-\frac{\boldsymbol{\upsilon}\cdot\boldsymbol{u}}{c^2}}\frac{\boldsymbol{\upsilon}}{c^2}\left(\frac{\boldsymbol{\upsilon}}{c}-\frac{\boldsymbol{u}}{c}\right)-\frac{\boldsymbol{u}\cdot\boldsymbol{\beta}}{c}\right)$$

$$\boldsymbol{P}_\delta\cdot\boldsymbol{R} = -i\hbar\frac{\partial\Phi}{\partial\delta}\left(\frac{\frac{\boldsymbol{u}}{c}\boldsymbol{\beta}}{1-\frac{\boldsymbol{\upsilon}\cdot\boldsymbol{u}}{c^2}}\left(1-\frac{\boldsymbol{\upsilon}\cdot\boldsymbol{u}}{c^2}\right)-\frac{\boldsymbol{u}\cdot\boldsymbol{\beta}}{c}\right)$$

$$\boldsymbol{P}_\delta\cdot\boldsymbol{R} = 0. \tag{321}$$

Similarly, from equation (109) and the second of equations (310) we get

$$\boldsymbol{P}_\omega\cdot\boldsymbol{R} = 0. \tag{322}$$

Both the momentum $\boldsymbol{P}_\delta$, and the momentum $\boldsymbol{P}_\omega$, are vertical to the vector $\boldsymbol{R}$ of Figure 6.

We will now prove that:

"The generalized photon carries intrinsic angular momentum $\boldsymbol{S}$, independent of the distance r. The component $\boldsymbol{S}_u$ of the intrinsic angular momentum $\boldsymbol{S}$ along the direction of the motion of the material particle does not depend upon the velocity $\boldsymbol{u}$ of the motion".

In Figure 4, the angular momentum $\boldsymbol{S}$ of the generalized photon with respect to the (constant) point of emission $E\left(x_p(w), y_p(w), z_p(w), w\right)$ is

$$\boldsymbol{S} = \boldsymbol{r} \times \boldsymbol{P}$$

and with equation (6) written in the form

$$\boldsymbol{r} = \frac{r}{c}\boldsymbol{\upsilon}$$

we get

$$\boldsymbol{S} = \frac{r}{c}\boldsymbol{\upsilon} \times \boldsymbol{P} = \frac{r}{c}\boldsymbol{\upsilon} \times (\boldsymbol{P}_i + \boldsymbol{P}_w + \boldsymbol{P}_\delta + \boldsymbol{P}_\omega). \tag{323}$$

Denoting

$$\begin{aligned} \boldsymbol{S}_i &= \frac{r}{c}\boldsymbol{\upsilon} \times \boldsymbol{P}_i \\ \boldsymbol{S}_w &= \frac{r}{c}\boldsymbol{\upsilon} \times \boldsymbol{P}_w \\ \boldsymbol{S}_\delta &= \frac{r}{c}\boldsymbol{\upsilon} \times \boldsymbol{P}_\delta \\ \boldsymbol{S}_\omega &= \frac{r}{c}\boldsymbol{\upsilon} \times \boldsymbol{P}_\omega \end{aligned} \tag{324}$$

equation (323) can be written as

$$\boldsymbol{S} = \boldsymbol{S}_i + \boldsymbol{S}_w + \boldsymbol{S}_\delta + \boldsymbol{S}_\omega. \tag{325}$$

From the first of equations (324) we have

$$\boldsymbol{S}_i = \frac{r}{c}\boldsymbol{\upsilon} \times \boldsymbol{P}_i$$

and with the second of equations (304) we get

$$\boldsymbol{S}_i = 0. \tag{326}$$

From the second of equations (324) we have

$$\boldsymbol{S}_w = \frac{r}{c}\boldsymbol{\upsilon} \times \boldsymbol{P}_w$$

and with the second of equations (308) we get

$$\boldsymbol{S}_w = \boldsymbol{0}. \tag{327}$$

From the third of equations (324) we have

$$\boldsymbol{S}_\delta = \frac{r}{c}\boldsymbol{\upsilon} \times \boldsymbol{P}_\delta$$

and with the second of equations (309) we have

$$\boldsymbol{S}_\delta = -i\hbar\frac{\partial\Phi}{\partial\delta}\frac{\upsilon}{c} \times \left(\frac{\boldsymbol{u}\cdot\boldsymbol{\beta}}{1 - \dfrac{\boldsymbol{\upsilon}\cdot\boldsymbol{u}}{c^2}}\frac{\boldsymbol{\upsilon}}{c^2} + \boldsymbol{\beta}\right)$$

$$\boldsymbol{S}_\delta = -i\hbar\frac{\partial\Phi}{\partial\delta}\frac{\boldsymbol{\upsilon}}{c} \times \boldsymbol{\beta}$$

and since it is $\frac{\upsilon}{c} \times \boldsymbol{\beta} = \boldsymbol{\gamma}$, we get

$$\boldsymbol{S}_\delta = -i\hbar\frac{\partial\Phi}{\partial\delta}\boldsymbol{\gamma}. \tag{328}$$

From the fourth of equations (324) we have

$$\boldsymbol{S}_\omega = \frac{r}{c}\boldsymbol{\upsilon} \times \boldsymbol{P}_\omega$$

and with the second of equations (310) we get

$$\boldsymbol{S}_\omega = -\frac{i\hbar}{\sin\delta}\frac{\partial\Phi}{\partial\omega}\frac{\boldsymbol{\upsilon}}{c} \times \left(\frac{\boldsymbol{u}\cdot\boldsymbol{\gamma}}{1 - \dfrac{\boldsymbol{\upsilon}\cdot\boldsymbol{u}}{c^2}}\frac{\boldsymbol{\upsilon}}{c^2} + \boldsymbol{\gamma}\right)$$

and since $\frac{\upsilon}{c} \times \boldsymbol{\gamma} = -\boldsymbol{\beta}$, we get

$$\boldsymbol{S}_\omega = \frac{i\hbar}{\sin\delta}\frac{\partial\Phi}{\partial\omega}\boldsymbol{\beta}. \tag{329}$$

Equation (325) can now be written as

$$\boldsymbol{S} = \frac{i\hbar}{\sin\delta}\frac{\partial\Phi}{\partial\omega}\boldsymbol{\beta} - i\hbar\frac{\partial\Phi}{\partial\delta}\boldsymbol{\gamma}. \tag{330}$$

We now calculate the component $\boldsymbol{S}_u$ of the angular momentum $\boldsymbol{S}$ along the direction of motion of the material particle.

For $\boldsymbol{u} \neq 0$ we have

$$\boldsymbol{S}_u = \frac{\boldsymbol{u}}{\|\boldsymbol{u}\|}\boldsymbol{S}$$

and with equation (330) we get

$$\boldsymbol{S}_u = \frac{\boldsymbol{u}}{\|\boldsymbol{u}\|}\left(\frac{i\hbar}{\sin\delta}\frac{\partial\Phi}{\partial\omega}\boldsymbol{\beta} - i\hbar\frac{\partial\Phi}{\partial\delta}\boldsymbol{\gamma}\right). \tag{331}$$

For constant-speed motion with velocity $\boldsymbol{u} = \begin{bmatrix} u \\ 0 \\ 0 \end{bmatrix}$, and taking into consideration equations (35) and (36), we obtain from equation (331)

$$S_u = \frac{i\hbar}{\sin\delta}\frac{\partial\Phi}{\partial\omega}(-\sin\delta)$$

$$S_u = -i\hbar\frac{\partial\Phi}{\partial\omega}. \tag{332}$$

In the case of constant-speed motion with velocity $\boldsymbol{u} = \begin{bmatrix} u \\ 0 \\ 0 \end{bmatrix}$, from the transformations of equations (124) $\omega' = \omega$, we conclude that the angular momentum S_u does not depend on the inertial reference frame. Furthermore, it does not depend on the angle δ, i.e. the angle formed between the direction of emission $\frac{\upsilon}{c}$ of the generalized photon and the velocity $\boldsymbol{u}$ of the material particle in Figure 6.

We will now study the changes in energy and momentum that take place during the motion of the generalized photon with velocity $\boldsymbol{\upsilon}$, after its emission by the material particle.

From the fundamental mathematical theorem, specifically from equation (86) for $f = E_i$, $f = E_w$, $f = E_\delta$ and $f = E_\omega$, we have

$$\begin{aligned}
&\frac{\partial E_i}{\partial t} + \boldsymbol{\upsilon}\cdot\nabla E_i = c\frac{\partial E_i}{\partial r}\\
&\frac{\partial E_w}{\partial t} + \boldsymbol{\upsilon}\cdot\nabla E_w = c\frac{\partial E_w}{\partial r}\\
&\frac{\partial E_\delta}{\partial t} + \boldsymbol{\upsilon}\cdot\nabla E_\delta = c\frac{\partial E_\delta}{\partial r}\\
&\frac{\partial E_\omega}{\partial t} + \boldsymbol{\upsilon}\cdot\nabla E_\omega = c\frac{\partial E_\omega}{\partial r}
\end{aligned}$$

and with the first of equations (304), (308), (309) and (310), we get

$$\begin{aligned}
&\frac{\partial E_i}{\partial t} + \boldsymbol{\upsilon}\cdot\nabla E_i = 0\\
&\frac{\partial E_w}{\partial t} + \boldsymbol{\upsilon}\cdot\nabla E_w = 0\\
&\frac{\partial E_\delta}{\partial t} + \boldsymbol{\upsilon}\cdot\nabla E_\delta = -\frac{c}{r}E_\delta\\
&\frac{\partial E_\omega}{\partial t} + \boldsymbol{\upsilon}\cdot\nabla E_\omega = -\frac{c}{r}E_\omega.
\end{aligned} \tag{333}$$

Similarly, after combining equations (87), (88), (89) with the second parts of equations (304), (308), (309) and (310), we get

$$\begin{aligned}
&\frac{\partial \boldsymbol{P}_i}{\partial t} + (\mathrm{grad}\ \boldsymbol{P}_\mathrm{i})\,\boldsymbol{\upsilon} = \boldsymbol{0}\\
&\frac{\partial \boldsymbol{P}_w}{\partial t} + (\mathrm{grad}\ \boldsymbol{P}_\mathrm{w})\,\boldsymbol{\upsilon} = \boldsymbol{0}\\
&\frac{\partial \boldsymbol{P}_\delta}{\partial t} + (\mathrm{grad}\ \boldsymbol{P}_\delta)\,\boldsymbol{\upsilon} = -\frac{c}{r}\boldsymbol{P}_\delta\\
&\frac{\partial \boldsymbol{P}_\omega}{\partial t} + (\mathrm{grad}\ \boldsymbol{P}_\omega)\,\boldsymbol{\upsilon} = -\frac{c}{r}\boldsymbol{P}_\omega.
\end{aligned} \tag{334}$$

From the equations of this paragraph we conclude that there are physical quantities that do not depend on the distance r. Such physical quantities are the energy-momentum pairs $(E_i, \boldsymbol{P}_i)$ and $(E_w, \boldsymbol{P}_w)$, as well as the angular momenta $\boldsymbol{S}$ and S_u. These quantities are defined for $r = 0$, that is, on the material particle. On the contrary, the energy-momentum pairs $(E_\delta, \boldsymbol{P}_\delta)$ and $(E_\omega, \boldsymbol{P}_\omega)$, as well as the rest energies $\frac{c\hbar}{r}\frac{\partial\Phi}{\partial\delta}$ and $\frac{c\hbar}{r\sin\delta}\frac{\partial\Phi}{\partial\omega}$, are defined only in the surrounding spacetime of the material particle, due to the appearance of the factor $\frac{1}{r}$. Furthermore, they vanish for $r \to +\infty$, while they attain large values for small values of r, i.e. close to the material particle.

5.7 The simplest case of a generalized photon

The simplest generalized photon arises in the case where the percentage Φ is constant:

$$\begin{aligned}
&\frac{\partial\Phi}{\partial t} = 0\\
&\nabla\Phi = 0.
\end{aligned} \tag{335}$$

In this case, equations (299) and (300) are rewritten, respectively

$$\begin{aligned}
E &= \Phi\frac{i\hbar}{1-\dfrac{\boldsymbol{\upsilon}\cdot\boldsymbol{u}}{c^2}}\frac{\partial m_0}{m_0\partial w}\\
\boldsymbol{P} &= \Phi\frac{i\hbar}{1-\dfrac{\boldsymbol{\upsilon}\cdot\boldsymbol{u}}{c^2}}\frac{\partial m_0}{m_0 c\partial w}\frac{\boldsymbol{\upsilon}}{c}
\end{aligned} \tag{336}$$

$$\begin{aligned}
E &= \Phi\frac{i\hbar}{1-\dfrac{\boldsymbol{\upsilon}\cdot\boldsymbol{u}}{c^2}}\frac{\partial q}{q\partial w}\\
\boldsymbol{P} &= \Phi\frac{i\hbar}{1-\dfrac{\boldsymbol{\upsilon}\cdot\boldsymbol{u}}{c^2}}\frac{\partial q}{qc\partial w}\frac{\boldsymbol{\upsilon}}{c}.
\end{aligned} \tag{337}$$

From the second of equations (335) we obtain

$$\nabla\Phi = 0$$

and from equation (301) we get

$$\frac{\partial\Phi}{\partial w}\nabla w + \frac{\partial\Phi}{\partial\delta}\nabla\delta + \frac{\partial\Phi}{\partial\omega}\nabla\omega = 0$$

and from the linear independence of the vectors $\nabla w, \nabla\delta, \nabla\omega$ (paragraph 2.5) we get

$$\begin{aligned}
&\frac{\partial\Phi}{\partial w} = 0\\
&\frac{\partial\Phi}{\partial\delta} = 0\\
&\frac{\partial\Phi}{\partial\omega} = 0.
\end{aligned} \tag{338}$$

Replacing equations (338) into the equations of the last paragraph causes the energy-momentum pairs

$$(E_w, \boldsymbol{P}_w), (E_\delta, \boldsymbol{P}_\delta), (E_\omega, \boldsymbol{P}_\omega)$$

to become zero, the angular momentum $\boldsymbol{S}$ becomes zero, and so do the rest energies $\frac{c\hbar}{r}\frac{\partial\Phi}{\partial\delta}$ and $\frac{c\hbar}{r\sin\delta}\frac{\partial\Phi}{\partial\omega}$. The energy-momentum pair $(E_i, \boldsymbol{P}_i)$, as given by equations (336), does not become zero. Therefore, the generalized photon is defined for $r = 0$, i.e. on the material particle.

We shall now prove that the interaction of the material particle with every generalized photon is instantaneous during the moment w of the emission of the generalized photon. More specifically, we shall prove that the generalized photon keeps its energy E and moment $\boldsymbol{P}$ constant, after its emission by the material particle.

From equation (86) of the fundamental mathematical theorem, and for $f = E$, we have

$$\frac{\partial E}{\partial t} + \upsilon\cdot\nabla E = c\frac{\partial E}{\partial r}. \tag{339}$$

From the first of equations (336), and since it holds that $m_0 = m_0(w)$, we get

$$\frac{\partial E}{\partial r} = 0 \tag{340}$$

and from equation (339) we see that

$$\frac{\partial E}{\partial t} + \boldsymbol{\upsilon} \cdot \nabla E = 0. \tag{341}$$

From equation (341) we conclude that the energy E of the generalized photon remains constant during its motion with velocity υ, after its emission by the material particle.

Combining equations (336) we obtain relation

$$\boldsymbol{P} = E\frac{\boldsymbol{\upsilon}}{c^2} \tag{342}$$

between the momentum $\boldsymbol{P}$ and energy E of the generalized photon.

From equation (87) for $f = \frac{E}{c}$, we obtain

$$\frac{\partial}{\partial t}\left(E\frac{\boldsymbol{\upsilon}}{c^2}\right) + \left(\text{grad}\left(\text{E}\frac{\boldsymbol{\upsilon}}{\text{c}^2}\right)\right)\boldsymbol{\upsilon} = \frac{\boldsymbol{\upsilon}}{c}\frac{\partial E}{\partial r}$$

and with equations (340) and (342) we get

$$\frac{\partial \boldsymbol{P}}{\partial t} + (\text{grad}\ \boldsymbol{P})\,\boldsymbol{\upsilon} = \boldsymbol{0}. \tag{343}$$

From equation (343) we conclude that the momentum $\boldsymbol{P}$ of the generalized photon remains constant during its motion with velocity $\boldsymbol{\upsilon}$, after its emission by the material particle.

According to equations (341) and (343), the generalized photon does not exchange energy and momentum with the material particle after its emission. The interaction between the material particle and every generalized photon takes place instantaneously at the moment of emission of the generalized photon. Furthermore, according to equation (342), there is a continuous flow of generalized photons moving with velocity $\boldsymbol{\upsilon}$, from the material particle into the surrounding spacetime, on the condition, of course, that the percentage Φ remains constant.

We can undertake a similar study for the generalized photon resulting from the selfvariation of the electric charge. It suffices to replace equations (336) with equations (337) in the above study.

5.8 The cosmological data "condensed" into a single equation

In the inertial frame of reference S', where the material particle is at rest, the first of equations (350) can be written as

$$E' = \Phi i\hbar\frac{\partial m_0}{m_0 \partial w'} + i\hbar\frac{\partial \Phi}{\partial t'}. \tag{344}$$

Summing in all directions of emission of generalized photons, and taking into consideration that $\sum E' = E_0$ and $\sum \Phi = 1$, from equation (344) we obtain

$$E_0 = i\hbar\frac{\partial m_0}{m_0 \partial w'}. \tag{345}$$

During the emission of the generalized photons by the material particle it is $r' = 0$, and equation (3) can be written as $w' = t'$, therefore we get $\frac{\partial m_0}{\partial w'} = \frac{dm_0}{dw'} = \frac{dm_0}{dt'} = \dot{m}_0$, and equation (345) can be written as

$$E_0 = i\hbar\frac{\dot{m}_0}{m_0} \tag{346}$$

which is equation (268).

In the inertial reference frame S', where the material particle is at rest, and for $r' = 0$, hence for $w' = t'$, the first of equations (282) can be written as

$$\dot{E}_0 = \frac{i}{\hbar} m_0 c^2 E_0. \tag{347}$$

Eliminating the rest energy E_0, we get

$$\begin{aligned} \left(i\hbar\frac{\dot{m}_0}{m_0}\right)^{\bullet} &= \tfrac{i}{\hbar} m_0 c^2 i\hbar\frac{\dot{m}_0}{m_0} \\ \left(i\hbar\frac{\dot{m}_0}{m_0}\right)^{\bullet} &= -\dot{m}_0 c^2 \\ \left(i\hbar\frac{\dot{m}_0}{m_0} + m_0 c^2\right)^{\bullet} &= 0 \end{aligned} \tag{348}$$

which is equation (270).

In paragraph 5.2 we derived equation (270) by combining equation (346) with the principle of conservation of energy. In the derivation we conducted in this paragraph we combined equation (346) with the symmetric law (282). Furthermore, from the derivation procedure we have followed, it becomes obvious that the percentage-function Φ does not play any role in equation (348), i.e. in equation (270).

If we borrow equation (394), $E_0 = i\hbar H$, from paragraph 7, and combine it with equation (346), we obtain $\frac{\dot{m}_0}{m_0} = H \sim 2 \times 10^{-18} s^{-1}$. In the cosmological data we observe the consequences of the real increase of the rest masses of the material particles, which takes place at an extremely slow rate.

In paragraph 7 the differential equation (348) is solved. As we shall see, this equation contains as information the totality of the cosmological data. The cosmological data are "condensed" within a single equation.

5.9 The generalized particle

From the previous study it becomes evident that the selfvariations correlate every material particle with the surrounding spacetime. Fundamental physical characteristics of the material particle, like the rest mass and the electric charge, are correlated with spacetime. Furthermore, each material particle contributes to the energy content of spacetime in a strictly defined manner.

The relation between the material particle and the surrounding spacetime is determined by two fundamental physical objects predicted by the theory of selfvariations: the generalized photon and the accompanying particle. These two

physical objects are related to each other since the accompanying particle results from the aggregation of the generalized photons. All the equations we have stated in the preceding paragraphs and preceding paragraphs, concern the relation of the material particle either with the generalized photon, or with the accompanying particle.

In the surrounding spacetime of the material particle, and for each generalized photon, we know exactly what is expressed by equation (263), $dS^2 = 0$: the generalized photon moves with velocity $\boldsymbol{\upsilon}$ of magnitude $\|\boldsymbol{\upsilon}\| = c$ in every inertial frame of reference. According to the second statement-axiom we have posed, equation $dS^2 = 0$ also holds for the accompanying particle, which, as an aggregation of generalized photons, is related with the propagation of the selfvariations in the four-dimensional spacetime. The question then arises, as to how equation $dS^2 = 0$ is expressed in the part of spacetime where the generalized photons aggregate.

The accompanying particle has rest energy E_0 and, therefore, rest mass $\frac{E_0}{c^2} \neq 0$. The combination $dS^2 = 0$ and $\frac{E_0}{c^2} \neq 0$ renders the accompanying particle an intermediate state between "matter" and the "photon". It is a completely new physical object predicted by the theory of selfvariations, which introduces us into an unknown territory of physical reality. The first question we have to answer is how do the relations $dS^2 = 0$ and $\frac{E_0}{c^2} \neq 0$ become compatible with each other.

About the intermediate state of matter we can give the following interpretation:

The aggregation of the generalized photons implies the co-incidence of different points ($dS^2 = 0$) in the part of spacetime where the aggregation takes place. This interpretation is in agreement with the strict application of the axioms of the theory of selfvariations.

At this point we are required to make two observations about the relation of the theory of selfvariations with the theory of relativity. These observations have to do with the relation between the energy content and the properties of spacetime.

For the derivation of the Lorentz-Einstein transformations we consider two observers who exchange signals moving with velocity c. If we consider the exchange of signals moving with a different velocity, for example acoustic signals, we end up with different transformations. Judging by the result, both on theoretical, and on experimental grounds, we know that the transformations derived by the first method are correct, whereas the transformations derived by the second method are wrong.

The theory of selfvariations predicts the generalized photon in the surrounding spacetime of the material particles. There is a continuous exchange of generalized photons between the material particles, in other words, a continuous exchange of signals moving with velocity c. The exchange of signals with velocity c is not simply a hypothesis we can make for the derivation of the Lorentz-Einstein transformations, but a continuous physical reality. Therefore, the theory of selfvariations strengthens the theoretical background of the special theory of relativity.

The general theory of relativity correlates the properties of spacetime with its energy content. The theory of selfvariations gives us the detailed contribution of each material particle to the energy content of spacetime. In the part of spacetime where the aggregation of generalized photons takes place, the material particle interacts with the accompanying particle. This interaction concerns a strictly distinct subset of the total energy content of spacetime. While we assume a unified spacetime, whose properties are defined by its total energy content, each particle interacts and is correlated with only a subset of the energy content of spacetime. In reality, every material particle occupies its "own" spacetime. For every material particle the properties of its "own" spacetime are determined by the generalized photons with which it interacts. Therefore, the co-incidence of different points of spacetime concerns the accompanying particle for every material particle, and does not constitute a general property of spacetime.

The law of selfvariations has been stated based on the accompanying particle. Relation (264), in combination with the symmetric laws (282) and (290), expresses the continuous interaction of the rest mass m_0 and the electric charge q of the material particle with the energy E_0 of the accompanying particle. Therefore, we cannot refer just to the material particle, or just to the accompanying particle. What exists in nature is the system of the two particles, which behaves as a "generalized particle" that occupies a part of spacetime.

The co-incidence of different points in the part of spacetime occupied by the generalized particle alters the trajectories and velocities of the generalized photons compared to the strictly defined trajectories and velocities we studied in the preceding paragraphs. In the case of co-incidence of all points belonging to this part of spacetime, the concepts of trajectory and velocity of the generalized photons loose their meaning. The trajectory and velocity of the material particle will suffer the same consequences, if the material particle belongs to the part of spacetime where the aggregation of the generalized photons takes place.

In Figures 4 and 6 imagine that, for the material particle, the points of spacetime within the interior of a sphere of centre E and radius r coincide. The physical object in the interior of the sphere constitutes a generalized particle with a specific rest mass. In every point of the spherical surface, the generalized photon moves with velocity υ of magnitude $\|\upsilon\| = c$. None of the axioms of special relativity and of the theory of selfvariations are violated. Furthermore, the co-incidence of different points of spacetime within the interior of the sphere, concerns the material particle, and does not constitute a general property of spacetime.

The investigation of the internal structure and physical

properties of the generalized particle is the central issue for the theory of selfvariations. We have to answer specific questions regarding the generalized particle, and develop specific methods for the study of its physical properties.

A fundamental question concerns the distribution of the total rest mass M_0 of the generalized particle, between the material particle (m_0) and the accompanying particle $\left(\frac{E_0}{c^2}\right)$. Of equal importance is the size of the portion of spacetime occupied by the generalized particle.

A basic method for the study of the generalized particle is the elimination of the velocity, which also represents the trajectory, from the equations of the theory of selfvariations. It is not the only method, though. In the following paragraph we present the basic study for the generalized particle.

6 The quantum phenomena as a consequence of the selfvariations

6.1 Introduction

The intermediate state between "matter" and "photon" predicted by the theory of selfvariations, is responsible for the quantum phenomena. The study of the generalized photon leads to the Schrödinger and the Klein-Gordon equations, as well as to the wave equation of Maxwell's theory of electromagnetism.

The elimination of the kinematic characteristics of the material particle from the equations of the selfvariations, emerges as the fundamental method for the study of the generalized particle and, eventually, of quantum phenomena. This is what is actually done by all the theories developed during the last century in order to interpret quantum phenomena.

The basic method for the study of the generalized particle is complemented by the percentage-function Φ. The Φ function has to do with the generalized photon and, by extension, with the generalized particle. Furthermore, it is related with the interactions of the material particles. Function Φ inextricably links the quantum phenomena with the interactions of the material particles. The investigation of its properties furthers the theory of selfvariations beyond the bounds of the present edition.

6.2 The distribution functions of the rest mass

According to equation (284)

$$M_0 = m_0 + \frac{E_0}{c^2} \tag{349}$$

the rest mass M_0 of the generalized particle is equal to the sum of the rest masses of the material particle (m_0) and the accompanying particle $\left(\frac{E_0}{c^2}\right)$. One way of studying the inner structure of the generalized particle is to study how the rest mass M_0 is distributed to each of the two particles. Knowing the sum of the rest masses m_0 and $\frac{E_0}{c^2}$, it suffices to calculate one of the "distribution functions", that is, one of the complex numbers $X = \frac{m_0}{M_0}$, $\Psi = \frac{E_0}{M_0c^2}$, $Z = \frac{m_0c^2}{E_0}$.

But it is

$$X + \Psi = \frac{m_0c^2}{M_0c^2} + \frac{E_0}{M_0c^2} = \frac{m_0c^2 + E_0}{M_0c^2}$$

and with equation (349) we get $X + \Psi = 1$. Therefore, it suffices to study either function Ψ

$$\Psi = \frac{E_0}{M_0c^2} \tag{350}$$

or function Z

$$Z = \frac{m_0c^2}{E_0} \tag{351}$$

in order to determine the distribution of the rest mass M_0 into m_0 and $\frac{E_0}{c^2}$.

Initially, we will study the effects of the selfvariations on the function Z. From equation (351) we have

$$\frac{\partial Z}{\partial t} = \frac{1}{E_0}\frac{\partial m_0c^2}{\partial t} - \frac{m_0c^2}{E_0^2}\frac{\partial E_0}{\partial t}$$

and with the firsts of equations (265) and (282) we get

$$\frac{\partial Z}{\partial t} = -\frac{1}{E_0}\frac{i}{\hbar}E_s m_0c^2 - \frac{m_0c^2}{E_0^2}\frac{i}{\hbar}\gamma m_0c^2E_0$$

and with equation (279) we get

$$\frac{\partial Z}{\partial t} = -\frac{1}{E_0}\frac{i}{\hbar}\gamma E_0 m_0c^2 - \frac{m_0c^2}{E_0^2}\frac{i}{\hbar}\gamma m_0c^2E_0$$

$$\frac{\partial Z}{\partial t} = -\frac{i}{\hbar}\frac{m_0c^2}{E_0}\gamma\left(m_0c^2 + E_0\right)$$

and with equation (349) we get

$$\frac{\partial Z}{\partial t} = -\frac{i}{\hbar}\frac{m_0c^2}{E_0}\gamma M_0c^2$$

and with equation (351)

$$\frac{\partial Z}{\partial t} = -\frac{i}{\hbar}\gamma M_0c^2Z. \tag{352}$$

From equation (351) we obtain

$$\nabla Z = \frac{1}{E_0}\nabla m_0c^2 - \frac{m_0c^2}{E_0^2}\nabla E_0$$

and with the second of equations (265) and also (266) we get

$$\nabla Z = \frac{1}{E_0}\frac{i}{\hbar}\boldsymbol{P}_s m_0c^2 + \frac{m_0c^2}{E_0^2}\frac{i}{\hbar}\gamma m_0\boldsymbol{u}E_0$$

and with equation (276) we have

$$\nabla Z = \frac{1}{E_0}\frac{i}{\hbar}E_s\frac{\boldsymbol{u}}{c^2}m_0c^2 + \frac{m_0c^2}{E_0^2}\frac{i}{\hbar}\gamma m_0\boldsymbol{u}E_0.$$

Using equation (279) we get

$$\nabla Z = \frac{1}{E_0}\frac{i}{\hbar}\gamma E_0 \frac{\boldsymbol{u}}{c^2} m_0 c^2 + \frac{m_0 c^2}{E_0^2}\frac{i}{\hbar}\gamma m_0 \boldsymbol{u} E_0$$

$$\nabla Z = \frac{i}{\hbar}\frac{m_0 c^2}{E_0^2}\gamma\left(\frac{E_0}{c^2} + m_0\right)\boldsymbol{u}.$$

Through equation (349) we get

$$\nabla Z = \frac{i}{\hbar}\frac{m_0 c^2}{E_0^2}\gamma M_0 \boldsymbol{u}$$

and with equation (351) we get

$$\nabla Z = \frac{i}{\hbar}\gamma M_0 \boldsymbol{u} Z. \tag{353}$$

The differential equations (352) and (353) offer the advantage that the rest mass M_0 that appears on their second part, does not depend on the selfvariations. On the other hand, they also have a disadvantage. We do not know the additional conditions we have to introduce for the rest mass M_0 in order to solve the system of differential equations (352) and (353). These additional conditions are related to a more general investigation of the equations of the theory of selfvariations, which is not included in the present edition.

We shall now study how the selfvariations affect function Ψ. From equation (350) we have

$$\frac{\partial \Psi}{\partial t} = \frac{\partial}{\partial t}\left(\frac{E_0}{M_0 c^2}\right)$$

and with equation (285) we obtain

$$\frac{\partial \Psi}{\partial t} = \frac{1}{M_0 c^2}\frac{\partial E_0}{\partial t}$$

and with the first of equations (282) we get

$$\frac{\partial \Psi}{\partial t} = \frac{1}{M_0 c^2}\frac{i}{\hbar}\gamma m_0 c^2 E_0$$

and from equation (350) we get

$$\frac{\partial \Psi}{\partial t} = \frac{i}{\hbar}\gamma m_0 c^2 \Psi. \tag{354}$$

From equation (350) we have

$$\nabla \Psi = \nabla\left(\frac{E_0}{M_0 c^2}\right)$$

and with equation (286) we obtain

$$\nabla \Psi = \frac{1}{M_0 c^2}\nabla(E_0)$$

and using the second of equations (282) we get

$$\nabla \Psi = \frac{1}{M_0 c^2}\left(-\frac{i}{\hbar}\gamma m_0 \boldsymbol{u} E_0\right)$$

and with equation (350) we arrive at

$$\nabla \Psi = -\frac{i}{\hbar}\gamma m_0 \boldsymbol{u} \Psi. \tag{355}$$

The differential equations (354) and (355) have the advantage that the rest mass m_0 of the material particle appears in their second part. This fact allows us to introduce additional conditions in order to solve the system of differential equations (354) and (355). We present this study in the following two paragraphs.

The distribution functions determine the distribution of the rest mass of the generalized particle between the material particle and the accompanying particle. For every point $A(x, y, z, t)$ in the part of spacetime where the generalized particle can reside, these distribution functions acquire specific values. These values, in turn, define the values of the rest masses m_0 and $\frac{E_0}{c^2}$.

The behavior of the generalized particle can be influenced by any cause that interacts with the generalized particle in the part of spacetime it occupies. An external cause can redistribute the rest mass of the generalized particle, directing it either to the material particle, or to the accompanying particle. In the first case, the generalized particle will behave as a material particle with a well-defined trajectory, energy, etc. In the second case, the generalized particle will spread out in spacetime, while the consequences resulting from the aggregation of the generalized photons will be strengthened and intensified. We observe such a case in the double-slit experiment for the electron and for material particles in general (we assume that the reader is familiar with the double-slit experiment).

The study of the distribution functions is a fundamental goal in order to understand the behavior of the generalized particle.

6.3 The Schrödinger equation

From equation (354) we have

$$\frac{\partial^2 \Psi}{\partial t^2} = \frac{i}{\hbar}\gamma m_0 c^2 \frac{\partial \Psi}{\partial t} + \frac{i}{\hbar}\gamma c^2 \Psi \frac{\partial m_0}{\partial t}$$

and with equation (354) and the first of equations (265), we get

$$\frac{\partial^2 \Psi}{\partial t^2} = -\frac{\gamma^2 m_0^2 c^4}{\hbar^2}\Psi + \frac{i}{\hbar}\gamma c^2 \Psi\left(-\frac{i}{\hbar}E_s m_0\right)$$

and with equation (279) we get

$$\frac{\partial^2 \Psi}{\partial t^2} = -\frac{\gamma^2 m_0^2 c^4}{\hbar^2}\Psi + \frac{i}{\hbar}\gamma c^2 \Psi\left(-\frac{i}{\hbar}\gamma E_0 m_0\right)$$

$$\frac{\partial^2 \Psi}{\partial t^2} = -\frac{\gamma^2 m_0^2 c^4}{\hbar^2}\Psi + \frac{\gamma^2 m_0 c^2 E_0}{\hbar^2}\Psi$$

$$\frac{\partial^2 \Psi}{\partial t^2} = -\frac{\gamma^2 m_0 c^4}{\hbar^2}\left(m_0 - \frac{E_0}{c^2}\right)\Psi. \tag{356}$$

From equation (355) we have

$$\nabla^2\Psi = -\frac{i}{\hbar}\gamma m_0 \boldsymbol{u}\nabla\Psi - \frac{i}{\hbar}\gamma\Psi\boldsymbol{u}\nabla m_0$$

and with equation (355) together with the second of equations (265), we get

$$\nabla^2\Psi = -\frac{\gamma^2 m_0^2 u^2}{\hbar^2}\Psi - \frac{i}{\hbar}\gamma\Psi\boldsymbol{u}\left(\frac{i}{\hbar}E_s\frac{\boldsymbol{u}}{c^2}m_0\right)$$

and with equation (279) we get

$$\nabla^2\Psi = -\frac{\gamma^2 m_0^2 u^2}{\hbar^2}\Psi - \frac{i}{\hbar}\gamma\Psi\boldsymbol{u}\left(\frac{i}{\hbar}\gamma E_0\frac{\boldsymbol{u}}{c^2}m_0\right)$$

$$\nabla^2\Psi = -\frac{\gamma^2 m_0^2 u^2}{\hbar^2}\Psi + \frac{\gamma^2 m_0 E_0 u^2}{c^2\hbar^2}\Psi$$

$$\nabla^2\Psi = -\frac{\gamma^2 m_0 u^2}{\hbar^2}\left(m_0 - \frac{E_0}{c^2}\right)\Psi. \qquad (357)$$

We now consider the case where the rest mass M_0 is mainly distributed to the material particle. This happens when

$$\left\|\frac{E_0}{m_0c^2}\right\| << 1$$

or when

$$E_0 \to 0.$$

Under these conditions equation (357) can be written as

$$\nabla^2\Psi = -\frac{\gamma^2 m_0^2 u^2}{\hbar^2}\Psi. \qquad (358)$$

We will now eliminate the velocity u from equation (358), within the framework of the analysis we performed in paragraph 5.9 for the generalized particle. For small velocities u, it is $\gamma \sim 1$, and equation (358) can be written as

$$\nabla^2\Psi = -\frac{m_0^2 u^2}{\hbar^2}\Psi. \qquad (359)$$

Furthermore, denoting by ε the constant sum of the kinetic energy $\frac{1}{2}m_0u^2$ and the potential energy $U = U(x, y, z)$ of the material particle, we have

$$\frac{1}{2}m_0u^2 + U = \varepsilon$$

$$u^2 = \frac{2(\varepsilon - U)}{m_0}.$$

Replacing factor u^2 into equation (359) we obtain

$$\nabla^2\Psi = -\frac{2m_0(\varepsilon - U)}{\hbar^2}\Psi \qquad (360)$$

which is the time-independent Schrödinger wave-function.

From the initial conditions, $\left\|\frac{E_0}{m_0c^2}\right\| << 1$ or $E_0 \to 0$, we set, and from equation (349) we obtain $m_0 \to M_0$, therefore equation (360) can be written in the form

$$\nabla^2\Psi = -\frac{2M_0(\varepsilon - U)}{\hbar^2}\Psi. \qquad (361)$$

From the derivation process we have followed it becomes obvious that the Schrödinger equation only approximately describes the internal structure of the generalized particle.

6.4 The Klein-Gordon equation

The way in which we chose to eliminate the velocity from equation (358) had as a consequence the appearance of the potential energy U in Schrödinger's equation (360). We will now eliminate the velocity u from function Ψin a different manner. Combining equations (356) and (357), we obtain

$$\frac{\partial^2\Psi}{\partial t^2} - c^2\nabla^2\Psi =$$
$$-\frac{\gamma^2 m_0 c^4}{\hbar^2}\left(m_0 - \frac{E_0}{c^2}\right)\Psi + \frac{\gamma^2 m_0 c^2 u^2}{\hbar^2}\left(m_0 - \frac{E_0}{c^2}\right)\Psi$$

$$\frac{\partial^2\Psi}{\partial t^2} - c^2\nabla^2\Psi = -\frac{\gamma^2 m_0 c^4}{\hbar^2}\left(1 - \frac{u^2}{c^2}\right)\left(m_0 - \frac{E_0}{c^2}\right)\Psi$$

and since $\gamma = \frac{1}{\sqrt{1-\frac{u^2}{c^2}}}$, we get

$$\frac{\partial^2\Psi}{\partial t^2} - c^2\nabla^2\Psi = -\frac{m_0 c^4}{\hbar^2}\left(m_0 - \frac{E_0}{c^2}\right)\Psi$$

$$\frac{\partial^2\Psi}{\partial t^2} - c^2\nabla^2\Psi + \frac{m_0 c^4}{\hbar^2}\left(m_0 - \frac{E_0}{c^2}\right)\Psi = 0. \qquad (362)$$

In the case where $\left\|\frac{E_0}{m_0c^2}\right\| << 1$ or $E_0 \to 0$, equation (362) can be written as

$$\frac{\partial^2\Psi}{\partial t^2} - c^2\nabla^2\Psi + \frac{m_0 c^4}{\hbar^2}\Psi = 0 \qquad (363)$$

which is the Klein-Gordon equation. With the conditions we posed, it follows that $m_0 \to M_0$ in equation (363).

Of particular interest is the case $m_0 = 0$, where from equation (362) we obtain

$$\frac{\partial^2\Psi}{\partial t^2} - c^2\nabla^2\Psi = 0$$

$$\nabla^2\Psi - \frac{\partial^2\Psi}{c^2\partial t^2} = 0. \qquad (364)$$

From equation (349) for $m_0 = 0$ we get $E_0 = M_0c^2$. Therefore, all of the rest energy of the generalized particle has shifted to the accompanying particle. Furthermore, we get $\|\Psi\| = \left\|\frac{E_0}{M_0c^2}\right\| = 1$. In every case we solve the differential equation (364), we should modify the final solution such that the wave-like behavior of a scalar quantity Ψ appears, for which we demand that $\|\Psi\| = 1$.

6.5 The central role of the percentage function Φ in the internal structure and the physical properties of the generalized particle

According to equations (302) and (303) the energy E and the momentum $\boldsymbol{P}$ of a single generalized photon depends on the percentage function Φ. Furthermore, according to equation (330), the intrinsic angular momentum $\boldsymbol{S}$ of a single generalized particle depends exclusively on the percentage function Φ. The generalized particle emerges in the part of spacetime where the aggregation of the generalized photons takes place. Therefore, the percentage function Φ plays a fundamental role, both for the internal structure, as well as for the physical properties of the generalized particle.

Function Φ allows the comprehension of the extent of the portion of spacetime occupied by the generalized particle. In paragraph 5.6 we determined the physical quantities that can only be defined in the surrounding spacetime of the material particle. These physical quantities are inversely proportional to the distance r. Therefore, the space occupied by the generalized photon can extend to infinity, with the consequences, of course, predicted by the corresponding equations for its energy, momentum, and angular momentum. Since each generalized photon can extend to infinity, the same also holds for the part of space where the aggregation of the generalized photons takes place. Therefore, the generalized particle can extend to infinity.

In the case of the simplest generalized photon, as we studied it in paragraph 5.7, there results an instantaneous interaction of the material particle with the accompanying particle. This interaction takes place at the instant of emission of the generalized photon, exactly at the point where the material point particle resides. Therefore, in this case the generalized particle is a point particle.

In conclusion, we can say that the generalized particle can extend from a point of spacetime up to an infinite distance from the material particle. Furthermore, in each case, the extent of the part of spacetime in which the generalized particle extends, is determined by the percentage function Φ.

For the derivation of the Schrödinger and the Klein-Gordon equations, we based our investigation on equation (349), $M_0 = m_0 + \frac{E_0}{c^2}$. A fundamental piece of information, related with the function Φ, is missing from this equation. The generalized photon carries rest energy, according to equations (314) and (315), which depends on the function Φ and the distance r. In other words, right from the start, the generalized photon, and therefore the generalized particle, are correlated with a rest energy in the surrounding spacetime of the material particle. The rest mass corresponding to this rest energy does not appear in equation (349). For the same reason, the angular momentum does not appear in the Schrödinger and the Klein-Gordon equations, since the internal angular momentum of the generalized photon depends exclusively on function Φ, according to equation (330).

Function Φ expresses the potential of a material particle to emit generalized photons of different energies for different directions. Theoretically, we cannot predict exactly how function Φ depends on the internal structure of the material particle. Quite likely we can do this by performing some measurements. But we can predict theoretically an important factor on which function Φ depends, that results from the continuous exchange of generalized photons between material particles. This exchange of generalized photons is equivalent to a variation of function Φ. According to equations (302), (303) and (330), the energy, momentum and intrinsic angular momentum of the generalized photon are exactly correlated with the variation of function Φ. We, therefore, come to the conclusion that the quantum phenomena are interrelated with the interactions of the material particles, the connecting link being function Φ. Function Φ is related with the interactions between material particles, but also with the energy of the generalized photons and, by extension, with the generalized particle.

In paragraph 5.9 we referred to the fundamental method for studying the generalized particle. We analyzed the reasons for which we have to expunge the velocity from the equations of the theory of selfvariations in order to study the internal structure and the physical properties of the generalized particle. Of equal importance is the inclusion of function Φ in the study of the generalized particle.

Observing the Schrödinger operators [22–26], as used in quantum mechanics, we realize that the first consequence of their use is the elimination of the kinematic characteristics of the material particle from the resulting differential equations. Function Φ does not appear in the final equations, since it does not exist as a concept within the physical theories of the last century. It is represented, though, by the physical quantities related with the interactions in which the material particle participates, by the potential energy or the generalized momentum of the material particle. Analogous is the procedure followed by Dirac [27] for the derivation of his eponymous equation.

One of the questions about the generalized particle, to which we deliberately did not refer in paragraph 5.9, is the probability of finding the material particle at a specific moment, in a specific position in the part of spacetime occupied by the generalized particle. There are many physical quantities related with the Schrödinger operators. Judging by the success of quantum mechanics, one way to study the generalized particle is through statistical interpretation. We must not forget, though, that a single cause suffices in order to shift the rest energy of the generalized particle, either towards the material particle, or towards the accompanying particle. One and only cause is sufficient for the corpuscular or wave-like behavior of the generalized particle to emerge.

By investigating the properties of function Φ or by making concrete hypotheses regarding function Φ, we can extend our study of quantum phenomena and the interactions of par-

ticles. On the contrary, in paragraph 5.8 we showed that equation (348) does not depend on function Φ. This allows us to solve it and investigate it completely. We present that study in the next paragraph.

7 The cosmological data as a consequence of the selfvariations

7.1 Introduction

The origin of matter is already recorded in the cosmological observational data. We just lacked a fundamental piece of information in order to decode it: the law of selfvariations.

The redshift of distant astronomical objects, the cosmic microwave background radiation and the information obtained by the analysis of this radiation, the increased luminosity distances of supernovae, the large-scale, as well as small-scale, structure of matter in the universe, the large-scale isotropy and flatness of the universe, the slight variation of the fine structure constant, and the arrow of time, all share the law of selfvariations as a common cause.

The law of selfvariations contains as information the entire corpus of the cosmological observational data, as we observe and record them since the time of Hubble. Behind the barrage of interventions made in order to bring the Standard Cosmological Model in agreement with the cosmological observational data, lies our ignorance about the fundamental law of selfvariations. The physical theories of the past century do not possess the necessary completeness in order to explain the cosmological observational data.

The improved scientific observation instruments we possess record persistently, and with ever increasing detail, the consequences of the law of selfvariations.

7.2 The fundamental equations

The cosmological data concern the observation of the Universe at long distances, that is, in the past. At a distant astronomical object, located at a distance r from Earth, the rest mass $m_0(r)$ of a material particle in the past is smaller, compared to the laboratory rest mass m_0 of the same material particle we measure "now" on Earth. The electric charge $q(r)$ also differs from the laboratory value q of the electric charge as measured "now" on Earth. We calculate the quantity $m_0(r)$ as a function of m_0, and $q(r)$ as a function of q. In this manner, we incorporate into our equations the consequences resulting from the internality of the Universe to the process of measurement.

In the following, and using the known physical laws, we determine the consequences of the selfvariations for distant astronomical objects. Furthermore, we can determine the consequences of the selfvariations in the electromagnetic spectra of the astronomical objects we receive "now" on Earth. We shall prove that equation (348)

$$\left(m_0c^2 + i\hbar\frac{\dot{m}_0}{m_0}\right)^{\bullet} = 0 \tag{365}$$

which holds for every material particle contains as information the entirety of the cosmological data.

We will solve equation (365) for a material particle in the case of a flat and static universe. This equation contains as information the redshift of distant astronomical objects. Furthermore, it predicts that the gravitational interaction cannot play the role attributed to it by the Standard Cosmological Model. It informs us that the gravitational interaction cannot lead the Universe either to collapse or to expansion. Consequently, there is no point of solving equation (365) within an expanding Universe.

Equation (365) contains as information the fact that the total energy of the Universe is zero. Therefore, after solving the equation, it can be verified a posteriori that the Universe is flat.

From equation (365) we have that

$$\left(m_0c^2 + i\hbar\frac{\dot{m}_0}{m_0}\right)^{\bullet} = 0$$

$$\left(\frac{i}{\hbar}m_0c^2 - \frac{\dot{m}_0}{m_0}\right)^{\bullet} = 0$$

$$\frac{i}{\hbar}m_0c^2 - \frac{\dot{m}_0}{m_0} = k. \tag{366}$$

Here, k is the constant of integration. From equation (366) we see that

$$m_0 = -\frac{ik\hbar}{c^2}\frac{1}{1-\exp(kt+\mu)}. \tag{367}$$

Here, μ is the constant of integration.

Let us suppose that we observe "now" on Earth, the electromagnetic spectrum of an astronomical object located at a distance r away from Earth. The emission of the electromagnetic spectrum from the astronomical object took place before a time interval $\Delta t = \frac{r}{c}$. According to equation (367) the rest mass $m_0(r)$ of the material particle at the moment of the emission of the corresponding electromagnetic spectrum was

$$m_0 = -\frac{ik\hbar}{c^2}\frac{1}{1-\exp\left(k\left(t-\frac{r}{c}\right)+\mu\right)}. \tag{368}$$

Combining equations (367) and (368) we have that

$$m_0(r) = m_0\frac{1-\exp(kt+\mu)}{1-\exp\left(k\left(t-\frac{r}{c}\right)+\mu\right)}.$$

Setting

$$A = \exp(kt+\mu) \tag{369}$$

we obtain

$$m_0(r) = m_0\frac{1-A}{1-A\exp\left(-\frac{kr}{c}\right)}. \tag{370}$$

Equation (370) expresses the rest mass $m_0(r)$ of the material particle in the distant astronomical object and before a

time interval $\Delta t = \frac{r}{c}$, compared with the laboratory value of the rest mass m_0 of the same material particle. In this way we include in the equations we state the consequences of the internality of the Universe with respect to the measurement process, as set forth in paragraph 4.9.

If we remove the imaginary unit i from equation (365), or replace it by any arbitrary constant $b \neq 0$, we will again end up with equations (369) and (370). The problems caused by the internality of the Universe with respect to the measurement procedure can only be evaded through equation (370). Only after comparing the rest masses $m_0(r)$ and m_0 can we measure the consequences of the selfvariations.

From equation (369) we obtain for the parameter A

$$\frac{dA}{dt} = \dot{A} = kA. \tag{371}$$

From equation (367) we also obtain

$$\dot{m}_0 = m_0 \frac{k \exp(kt + \mu)}{1 - \exp(kt + \mu)}.$$

Through equation (369) we see that

$$\dot{m}_0 = m_0 \frac{kA}{1 - A}. \tag{372}$$

Combining equations (268) and (372) we obtain

$$E_0 = i\hbar \frac{kA}{1 - A}. \tag{373}$$

In the case of the electric charge the corresponding equation to equation (365) is the second of equations (292)

$$\left(q + \frac{i\hbar}{V_0} \frac{\dot{q}}{q}\right)^{\bullet} = 0. \tag{374}$$

This gives us the corresponding solution

$$q(r) = q \frac{1 - B}{1 - B \exp\left(-\dfrac{k_1 r}{c}\right)} \tag{375}$$

$$B = \exp(k_1 t + \mu_1) \tag{376}$$

$$\frac{dB}{dt} = \dot{B} = k_1 B. \tag{377}$$

Here, k_1 and μ_1 are the constants of integration.

The corresponding equation to equation (372) is equation

$$\dot{q} = q \frac{k_1 B}{1 - B}. \tag{378}$$

Combining the first of equations (292) with equation (378) we obtain

$$q_i V_0 = i\hbar \frac{k_1 B}{1 - B}. \tag{379}$$

This equation is the corresponding equation to equation (373).

If we remove from equation (374) the imaginary unit i, or if we replace it by any arbitrary constant $b \neq 0$, we will still arrive at equations (375) and (376). The problems caused by the internality of the Universe with respect to the measurement procedure can only be evaded through equation (375). We can only measure the consequences of the selfvariations by comparing the electric charges $q(r)$ and q.

7.3 The redshift of the far distant astronomical objects

The wavelength λ of the linear spectrum of an atom is inversely proportional to the factor $m_0 q^4$, where m_0 is the rest mass and q the electric charge of the electron. We denote by λ the wavelength of the linear spectrum we observe "now" on Earth, and which originates from the atoms of an astronomical object located at distance r. With λ_0 we denote the wavelength of the same kind of atom as measured in the laboratory "now" on Earth.

We have that

$$\frac{\lambda}{\lambda_0} = \frac{m_0 q^4}{m_0(r) q^4(r)}.$$

Using equations (370) and (375) we obtain

$$\frac{\lambda}{\lambda_0} = \frac{1 - A \exp\left(-\dfrac{kr}{c}\right)}{1 - A} \left(\frac{1 - B \exp\left(-\dfrac{k_1 r}{c}\right)}{1 - B} \right)^4. \tag{380}$$

For the redshift z of the astronomical object we obtain

$$z = \frac{\lambda - \lambda_0}{\lambda_0}$$

$$z = \frac{\lambda}{\lambda_0} - 1.$$

Using equation (380) we see that

$$z = \frac{1 - A \exp\left(-\dfrac{kr}{c}\right)}{1 - A} \left(\frac{1 - B \exp\left(-\dfrac{k_1 r}{c}\right)}{1 - B} \right)^4 - 1. \tag{381}$$

This equation constitutes the full mathematical expression for the redshift z of the linear spectrum of distant astronomical objects.

We shall now perform an approximation. From the cosmological data we know that the fine structure constant

$$\alpha = \frac{e^2}{4\pi\varepsilon_0 c\hbar}$$

remains constant for observations we make at very large distances from Earth. Therefore, the value of the electric charge

$q(r)$ differs minimally from the laboratory value q in the region of the Universe we observe. Therefore, we can write equation (381) in a simpler form, that is

$$z = \frac{1 - A\exp\left(-\frac{kr}{c}\right)}{1 - A} - 1. \tag{382}$$

Here, we used the approximation $q(r) = q$.

Equation (382) holds for the regions of the Universe that can be surveyed by the scientific observation instruments we currently have at our disposal. We shall return to the issue of the fine structure constant in another paragraph.

From equation (369) we see that

$$A > 0. \tag{383}$$

According to equation (382), and since the value of the redshift z increases with the distance r, it holds that

$$k > 0. \tag{384}$$

From equation (382), and for $r \to +\infty$, we obtain

$$z_{\max} = \frac{1}{1 - A} - 1$$

$$z_{\max} = \frac{A}{1 - A}. \tag{385}$$

We have that $z_{\max} > 0, A > 0$, as given in relation (383), thus we get

$$1 - A > 0$$

$$A < 1. \tag{386}$$

Now, it holds that

$$z < z_{\max}.$$

Using equation (385) we obtain

$$z < \frac{A}{1 - A}.$$

Due to relation (386) we obtain

$$\begin{aligned} &z(1 - A) < A \\ &z - zA < A \\ &z < (1 + z)A \\ &\frac{z}{1 + z} < A. \end{aligned}$$

Through relation (386) we finally arrive at

$$\frac{z}{1 + z} < A < 1. \tag{387}$$

This inequality holds for every redshift z, and allows us to estimate the range of values the parameter A acquires.

From equation (382), and for $r = 0$, we obtain $z = 0$, thus

$$z' = \frac{dz}{dr} = \frac{1 - A\exp\left(-\frac{kr}{c}\right)}{1 - A} \frac{kA\exp\left(-\frac{kr}{c}\right)}{c(1 - A)}.$$

For $r = 0$ we get

$$z'(0) = \left.\frac{dz}{dr}\right|_{r=0} = \frac{kA}{c(1 - A)}.$$

We expand equation (382) giving $z = z(r)$ into a Taylor series, and only to first order terms

$$\begin{aligned} z(r) &= z(0) + z'(0) \\ z(r) &= 0 + \frac{kA}{c(1 - A)} r \\ cz &= \frac{kA}{1 - A} r. \end{aligned}$$

Comparing with Hubble's law $cz = Hr$, we obtain

$$\frac{kA}{1 - A} = H. \tag{388}$$

where H is the Hubble parameter.

From equation (388) we obtain $k = H\frac{1-A}{A}$. The range of values of parameter A, as determined from inequality (387), allows us to estimate the extremely small value of the constant k. Now, according to equation (371), the parameter A increases at an extremely slow rate, and remains practically constant in the measurements we conduct.

For the energy E, which results during nuclear fission, nuclear fusion, and more generally, every case where the conversion of rest mass into energy takes place, we obtain

$$\frac{E(r)}{E} = \frac{m_0(r)c^2}{m_0c^2}.$$

Using equation (370) we see that

$$\frac{E(r)}{E} = \frac{1 - A}{1 - A\exp\left(-\frac{kr}{c}\right)}$$

$$E(r) = E\frac{1 - A}{1 - A\exp\left(-\frac{kr}{c}\right)}. \tag{389}$$

For the photons which result from the conversion of mass into energy we have

$$\frac{\lambda_\gamma}{\lambda_{0\gamma}} = \frac{\frac{ch}{E(r)}}{\frac{ch}{E}} = \frac{E}{E(r)}.$$

Using equation (389) we obtain

$$\frac{\lambda_\gamma}{\lambda_{0\gamma}} = \frac{1 - A\exp\left(-\frac{kr}{c}\right)}{1 - A}$$

$$\frac{\lambda_\gamma - \lambda_{0\gamma}}{\lambda_{0\gamma}} = \frac{1 - A\exp\left(-\frac{kr}{c}\right)}{1 - A} - 1$$

$$z_\gamma = \frac{1 - A\exp\left(-\frac{kr}{c}\right)}{1 - A} - 1. \tag{390}$$

Equations (390) and (382) are identical. However, beyond the limits reached by our current observations, the redshift z of the linear spectrum is given in general by equation (381).

From equation (382) we obtain

$$1 + z = \frac{1 - A\exp\left(-\frac{kr}{c}\right)}{1 - A}. \tag{391}$$

Combining equations (370) and (391) we have that

$$m_0(z) = \frac{m_0}{1 + z}. \tag{392}$$

Combining equations (389) and (391) we see that

$$E(z) = \frac{E}{1 + z}. \tag{393}$$

Combining equations (373) and (388) we obtain

$$E_0 = i\hbar H \tag{394}$$

for the laboratory value of the energy E_0.

7.4 The graphs of the functions r = r(z) and R = R(z)

From equation (382) we have that

$$z = \frac{1 - A\exp\left(-\frac{kr}{c}\right)}{1 - A} - 1$$

$$z = \frac{A}{1 - A}\exp\left(-\frac{kr}{c}\right).$$

Solving for r we obtain

$$r = \frac{c}{k}\ln\left(\frac{A}{A - z(1 - A)}\right). \tag{395}$$

This equation gives the distance r of the astronomical object as a function of the redshift z.

From equation (388) we obtain $k = H\frac{1-A}{A}$, and after replacing the constant k into equation (395), we get

$$r = \frac{c}{H}\frac{A}{1 - A}\ln\left(\frac{A}{A - z(1 - A)}\right). \tag{396}$$

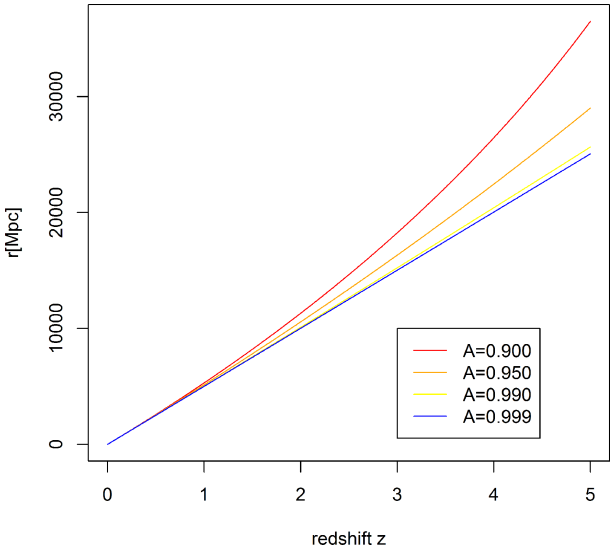

Fig. 10: The graph of distance r of a distant astronomical object as a function of the redshift z. As we increase the value of the parameter A from 0.900 to 0.999, the curve $r = r(z)$ approaches a straight line. The graph has been made with $H = 60\frac{\text{km}}{\text{sMpc}}$ as the value of Hubble's constant.

This equation is more convenient than equation (395), since we know the value of the Hubble parameter H, as well as the range of values of the parameter A from inequality (387), that is

$$\frac{z}{1 + z} < A < 1.$$

In Figure 10 we present the graph of the curve $r = r(z)$ for $H = 60\frac{\text{km}}{\text{sMpc}}$, and for the values of $A = 0.900$, $A = 0.950$, $A = 0.990$, $A = 0.999$ up to $z = 5$. We observe that as the value of the parameter A increases, the curve tends to be a straight line.

We shall now prove that for $A \to 1^-$ the equivalent equations (382) and (396) tend to Hubble's law

$$cz = Hr \tag{397}$$

From equation (388) we have $k = \frac{1-A}{A}H$, and after substituting into equation (382), we obtain

$$z = \frac{1 - A\exp\left(-\frac{1 - A}{A}\frac{H}{c}r\right)}{1 - A} - 1.$$

We denote by $x = \frac{1-A}{A}$, therefore $x \to 0^+$ for $A \to 1^-$, and

$A = \frac{1}{x+1}$, so we have

$$z = \frac{1-\frac{1}{x+1}\exp\left(-x\frac{Hr}{c}\right)}{1-\frac{1}{x+1}} - 1 = \frac{x+1-\exp\left(-x\frac{Hr}{c}\right)}{x} - 1$$

$$\lim_{A\to 1^-}(z) = \lim_{x\to 0^+}(z) = \lim_{x\to 0^+}\left(\frac{x+1-\exp\left(-x\frac{Hr}{c}\right)}{x} - 1\right) \overset{\frac{0}{0}}{=}$$

$$\lim_{x\to 0^+}\left(1+\frac{Hr}{c}\exp\left(-x\frac{Hr}{c}\right)-1\right) = \frac{Hr}{c}.$$

Equation (396) gives the distance r of the astronomical object, when we know the value of its redshift z. On the other hand, if we measure the distance based on the luminosity of the astronomical object, we shall always find it to be larger than the one given by equation (396). The reason is simple: The energy feeding the radiation of the astronomical objects originates from nuclear fusion, and more generally, from the conversion of rest mass into energy. According to equation (389), this energy $E(r)$ is less than the expected energy E. Therefore, the luminosity of the astronomical object is itself lower than the standard luminosity we use.

The luminosity distance R of an astronomical object is defined by equation

$$J = \frac{1}{4\pi R^2}\frac{dE}{dt}. \tag{398}$$

In this equation, J denotes the power per unit surface we receive from the astronomical object, whereas the power $\frac{dE}{dt}$ refers to the "standard candle" we are using.

If the real distance of the astronomical object is r, then we obtain for the power per unit surface J

$$J = \frac{1}{4\pi r^2}\frac{dE(r)}{dt}. \tag{399}$$

From equations (398) and (399) we get

$$\frac{1}{R^2}\frac{dE}{dt} = \frac{1}{r^2}\frac{dE(r)}{dt}.$$

Using equation (393) we have that

$$\frac{1}{R^2}\frac{dE}{dt} = \frac{1}{r^2}\frac{1}{1+z}\frac{dE}{dt}$$

$$R^2 = r^2(1+z)$$

$$R = r\sqrt{1+z}. \tag{400}$$

Combining equations (400) and (396) we see that

$$R = \frac{c}{H}\frac{A}{1-A}\sqrt{1+z}\ln\left(\frac{A}{A-z(1-A)}\right). \tag{401}$$

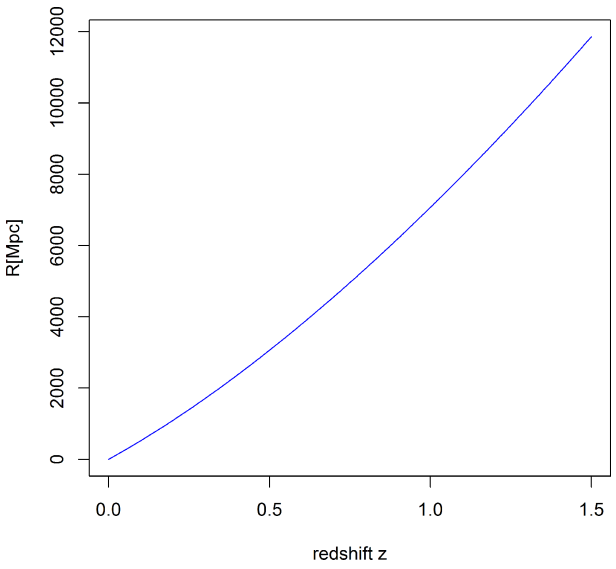

Fig. 11: The graph of the luminosity distance R of astronomical objects as a function of the redshift z. The measurement of the luminosity distances of type Ia supernova confirms the theoretical prediction of the law of selfvariations.

The measurements conducted for the determination of Hubble's constant H, have not taken into account the consequences of equation (400). Even for the case of small values of the redshift z it holds that $R > r$. The measurement of Hubble's parameter H with the use of the luminosity distances of astronomical objects is correct only for very small values of z, where it holds that $R \sim r$. Such measurements result in a value of $H = 60\frac{\text{km}}{\text{sMpc}}$. Measurements performed have included astronomical objects with large values of the redshift z, thus increasing the value of the parameter H to values ranging between 72 and $74\frac{\text{km}}{\text{sMpc}}$.

Today, we perform measurements with high accuracy. Taking into consideration the consequences of equation (400) we expect the parameter H to be measured close to $60\frac{\text{km}}{\text{sMpc}}$, independently of the redshift z of the astronomical object. We, of course, refer to measurements of the parameter H, on the basis of the luminosity distances of astronomical objects.

Equally well to equation (401) we can also use the equation which results after combining equations (400) and (397), that is

$$R = \frac{c}{H}z\sqrt{1+z}. \tag{402}$$

For $H = 60\frac{\text{km}}{\text{sMpc}}$ and $c = 3\times 10^5\frac{\text{km}}{\text{s}}$ this can be written as

$$R = 5000\,z\sqrt{1+z}. \tag{403}$$

The luminosity distance R is given in Mpc. In the graph of figure 11 we present the graph of the function $R = R(z)$, as given in equation (403) and up to values of the redshift $z = 1.5$.

Type *Ia* supernova are cosmological objects for which we can measure their luminosity distance R. Furthermore, this measurement can be conducted at large distances, where the consequences of equation (393) are measurable.

At the end of the last century this kind of measurements were conducted by independent scientific groups. The graph thet results from these measurements exactly matches graph 7.4.2 which is predicted theoretically by the theory of self-variations. In order to explain the inconsistency of the Standard Cosmological Model with graph 7.4.2, the existence of dark energy was invented and introduced [28–30].

7.5 Gravity cannot play the role attributed to it by the Standard Cosmological Model

From equation (388) we obtain $k = \frac{1-A}{A}H$, and

$$\frac{k}{c} = \frac{1-A}{A}\frac{H}{c}. \tag{404}$$

For $H = 60\frac{\text{km}}{\text{sMpc}}$, $A = 0.999$, $c = 3\times 10^5\frac{\text{km}}{\text{s}}$ we have that

$$\frac{k}{c} = 2\times 10^{-7}\frac{1}{\text{Mpc}}. \tag{405}$$

We replace this value of $\frac{k}{c}$ into equation (370) and obtain

$$m_0(r) = m_0\frac{0.001}{1-0.999\,\exp\left(-2\times 10^{-7}r\right)}. \tag{406}$$

Here, the distance r is measured in Mpc.

For values of r of the order of magnitude of kpc, equation (406) gives that $m_0(r) = m_0$. Therefore, the strength of the gravitational interaction is not affected in the scale of galactic distances. For example, for distance $r = 100$ kpc, equation (406) gives $m_0(r) = 0.99999\ m_0$. Therefore, the selfvariations do not affect the stability of the solar system, galaxies, and galaxy clusters.

On the contrary, for distances of order of magnitude of Mpc, equation (406) predicts a clearly smaller value of $m_0(r)$, compared to m_0. For example, for $r = 100$ Mpc equation (406) gives $m_0(r) = 0.98\ m_0$. The strength of the gravitational interaction exerted on our galaxy by a galaxy from a distance of 100 Mpc is 98% of the expected. For $r = 2\times 10^3$ Mpc equation (406) gives $m_0(r) = 0.7145\ m_0$. The strength of the gravitational interaction exerted by a galaxy, which is located at a distance of 2000 Mpc, on our galaxy is only 71.45% of the expected.

Therefore, we conclude that due to the selfvariations the gravitational interaction is weakened at cosmological distances and cannot play the role attributed to it by the Standard Cosmological Model. The gravitational interaction dominates and rules at a local level, at scales of a few hundreds or thousands of kpc.

We note that if we chose a different value for the parameter A, from the values permitted by inequality (387), all the arithmetic values appearing in equation (406) shall be altered. However, the same conclusions will be drawn about the relation between rest masses $m_0(r)$ and m_0.

The rest mass is given as a function of the redshift z from equation (392)

$$m_0(z) = \frac{m_0}{1+z}.$$

For $z = 0.1$ we get $m_0(z) = 0.9091\ m_0$, for $z = 1$ we have $m_0(z) = 0.5\ m_0$, and for $z = 9$ we see that $m_0(z) = 0.1\ m_0$. The strength of the gravitational interaction exerted by an astronomical object with redshift $z = 9$ on our galaxy is only 10% of the expected.

For even greater distances the gravitational interaction practically vanishes.

7.6 The very early Universe

All the equations we have stated in this paragraph are compatible with the condition $r \to \infty$. The equations are stated in such a way that the condition $r \to \infty$ offers us information about the very early Universe.

For $r \to \infty$ equation (370) gives

$$m_0(r \to \infty) \to m_0(1-A). \tag{407}$$

The inequality (387)

$$\frac{z}{1+z} < A < 1$$

holds for every value of the redshift z, hence $A \to 1$. Therefore, from relation (407) we conclude that the initial form of the Universe only slightly differs from the vacuum.

Similarly, from equation (375) we have that

$$q(r \to \infty) \to q(1-B). \tag{408}$$

This relation does not lead to the same consequences as relation (407). We know that the electric charge exists in opposite physical quantities in the Universe. Because of this, the total electric charge of the Universe is zero. Relation (407) informs as that the energy content of the very early Universe also tends to zero. The very early Universe only slightly differs from the vacuum. It possesses, though, a very important property which determines its evolution. It is temporally variable due to the selfvariations.

The increase of the rest masses and the electric charges destroys the initial homogeneity and state of rest, induces the first minute motions of the particles, and shifts the system to a temperature slightly above 0 K (temperature reflects the kinetic state of the particles in the system). The evolution of the selfvariations with the passage of time leads the Universe to the form in which we observe it today.

In general, this is the prediction for the begining and evolution of the Universe from the equations we have stated. This prediction is also verified from the calculations presented in the following paragraphs.

7.7 The Universe is flat

From the principle of conservation of energy we conclude that the total energy content of the Universe is constant, and remains the same at every moment. Relation (407) informs us that the energy content of the very early Universe tends to zero. Therefore, the same holds today as we observe the Universe. Because of this, the Universe is flat.

The difference between the current state of the Universe and its initial state is the following: The rest masses of particles have increased, but this increase is counterbalanced by the generalized photons that flood spacetime, and by the strengthening of all kinds of negative potential energies that result as a consequence.

The observations conducted by the COBE and WMAP satellites confirm that the Universe is flat. Other observational data lead us to the same conclusion.

7.8 The origin of the cosmic microwave background radiation

The laboratory value for the Thomson scattering coefficient [31, 32] is

$$\sigma_\tau = \frac{8\pi}{3}\frac{q^4}{m_0^2 c^2}. \tag{409}$$

Here, q and m_0 are the electric charge and the rest mass of the electron, respectively. At a distant astronomical object the Thomson coefficient is

$$\sigma_\tau(r) = \frac{8\pi}{3}\frac{q^4(r)}{m_0^2(r)\,c^2}. \tag{410}$$

Combining these equations we get that

$$\frac{\sigma_\tau(r)}{\sigma_\tau} = \left(\frac{m_0}{m_0(r)}\right)^2\left(\frac{q(r)}{q}\right)^4. \tag{411}$$

From the observations we have made on the variation of the fine structure constant we know that, for large distances r, it holds that $q(r) = q$. Therefore, at a very good approximation, equation (411) can be written as

$$\frac{\sigma_\tau(r)}{\sigma_\tau} = \left(\frac{m_0}{m_0(r)}\right)^2.$$

Using equation (370) we obtain that

$$\frac{\sigma_\tau(r)}{\sigma_\tau} = \left(\frac{1 - A\exp\left(-\dfrac{kr}{c}\right)}{1 - A}\right)^2. \tag{412}$$

For very large distances ($r \to \infty$) very close to the initial state of the Universe, and at a temperature of about 0 K, equation (412) gives

$$\frac{\sigma_\tau(r \to \infty)}{\sigma_\tau} = \left(\frac{1}{1 - A}\right)^2. \tag{413}$$

But according to inequality (387), $A \to 1$. Therefore, in the very distant past, and for a temperature of the Universe just slightly above 0 K, the Thomson scattering coefficient acquires enormous values, rendering the Universe opaque. The cosmic microwave background radiation we observe today, originates in this phase of the evolution of the Universe. The conditions we described refer to the whole expanse of the Universe. That is why the cosmic microwave background radiation seems to originate "from everywhere".

Equation (412) gives the value of the scattering coefficient at distant astronomical objects. Combining this equation with equation (382) gives

$$\frac{\sigma_\tau(z)}{\sigma_\tau} = (1 + z)^2$$

$$\sigma_\tau(z) = \sigma_\tau(1 + z)^2. \tag{414}$$

This equation is easier to use, since it expresses the Thomson scattering coefficient as a function of the redshift z of the distant astronomical object. We can also write equation (414) in the form

$$\sigma_\tau(z) = \frac{8\pi}{3}\frac{e^4}{m_e^2 c^2}(1 + z)^2 \tag{415}$$

where e and m_e denote the electric charge and the mass of the electron, respectively.

The Thomson coefficient concerns the scattering of photons of low energy E. For high energy photons it is replaced by the Klein-Nishina coefficient, given in the laboratory by

$$\sigma = \frac{3}{8}\sigma_\tau\frac{m_0 c^2}{E}\left[\ln\left(\frac{2E}{m_0 c^2}\right) + \frac{1}{2}\right] \tag{416}$$

and by relation

$$\sigma(z) = \frac{3}{8}\sigma_\tau(z)\,\frac{m_0(z)\,c^2}{E(z)}\left[\ln\left(\frac{2E(z)}{m_0(z)\,c^2}\right) + \frac{1}{2}\right] \tag{417}$$

for the distant astronomical object.

From equations (392) and (393) we obtain

$$\frac{m_0(z)\,c^2}{E(z)} = \frac{m_0 c^2}{E}.$$

Therefore, equation (417) can be written as

$$\sigma(z) = \frac{3}{8}\sigma_\tau(z)\,\frac{m_0 c^2}{E}\left[\ln\left(\frac{2E}{m_0 c^2}\right) + \frac{1}{2}\right].$$

Using equation (416) we have

$$\frac{\sigma}{\sigma(z)} = \frac{\sigma_\tau}{\sigma_\tau(z)}.$$

Using equation (414) we take that

$$\sigma(z) = \sigma(1 + z)^2. \tag{418}$$

The two scattering coefficients depend in the same way upon the redshift z, and the distance r.

7.9 The decrease of the atomic ionization energies at distant astronomical objects

The ionization and excitation energy X_n of the atoms is proportional to the factor $m_0 q^4$, where m_0 is the rest mass of the electron and q is its electric charge. Thus, we have

$$\frac{X_n(r)}{X_n} = \frac{m_0(r)}{m_0}\left(\frac{q(r)}{q}\right)^4.$$

After applying the familiar approximation $q(r) = q$ we obtain

$$\frac{X_n(r)}{X_n} = \frac{m_0(r)}{m_0}.$$

Using equation (370) we have

$$\frac{X_n(r)}{X_n} = \frac{1-A}{1-A\exp\left(-\dfrac{kr}{c}\right)}. \tag{419}$$

Through equation (382) we see that

$$X_n(z) = \frac{X_n}{1+z}. \tag{420}$$

According to this equation the redshift z affects the rate of ionization of the atoms in distant astronomical objects. Boltzmann's equation

$$\frac{N_n}{N_1} = \frac{g_n}{g_1}\exp\left(-\frac{X_n}{KT}\right) \tag{421}$$

expresses the number of the ionized atoms N_n occupying the energy level n in a stellar surface which is at thermodynamic equilibrium. With X_n we denote the excitation energy from the energy level 1 to the level n, T stands for the temperature of the stellar surface in $Kelvin$, $K = 1.38\times10^{-23}$ J/K is Boltzmann's constant, and g_n is the degree of degeneracy multiplicity of level n, that is, the number of energy levels into which level n splits in the presence of a magnetic field.

Combining equations (420) and (421), we obtain for the distant astronomical object relation

$$\frac{N_n}{N_1} = \frac{g_n}{g_1}\exp\left(-\frac{X_n}{KT(1+z)}\right). \tag{422}$$

In the case of the hydrogen atom, for $n = 2$, $X_2 = 10.15$ eV $= 16.24\times10^{-19}$ J, $g_1 = 2$, $g_2 = 8$ and for a solar surface temperature $T \sim 6000$ K, equation (421) shows that only one atom out of 10^8 occupies the $n = 2$ state. At the same temperature, equation (422) gives that for a redshift value of $z = 1$ we have $\frac{N_2}{N_1} = 2.2\times10^{-4}$, for $z = 2$ we have $\frac{N_2}{N_1} = 5.8\times10^{-3}$, and for $z = 5$ we have $\frac{N_2}{N_1} = 0.15$.

The conclusions drawn from the current and the previous paragraph demand a reexamination of the conclusions we have drawn from the observation of the electromagnetic spectrum of distant astronomical objects.

For very large distances, that is, in the very distant past, equation (419) gives

$$X_n(r\to\infty) = X_n(1-A). \tag{423}$$

This equation informs us that the very early Universe was ionized at some stage [33]. The ionization energies of the atoms had very small values. We can reach the same conclusion if we substitute into equation (420) very large values of the variable z, or if in equation (421) we replace the energy X_n with $X_n(1-A)$.

7.10 On the fine structure constant

In the preceding paragraphs we saw that due to the manifestation of the selfvariations, energy, momentum, angular momentum and electric charge flow from the material particles to the surrounding spacetime. The first consequence of the selfvariations is the potential to transfer energy, momentum, angular momentum and electric charge from one material particle to another, i.e. the interaction between the material particles. The gravitational and electromagnetic interactions determine the starting point for the quantitative determination of the selfvariations. Because of this, we supposed that the rest masses and the electric charges, and not any other physical quantity, vary with the passage of time. We offer this remark since, at cosmological scales, equation (365) justifies all of the cosmological observational data we possess, and it could be supposed that the electric charge remains constant. Such an assumption cannot hold within the framework of the theory of selfvariations, where the selfvariations of the electric charge are responsible for the electromagnetic field.

By analyzing the electromagnetic spectra reaching Earth from distant quasars from distances up to $6\times10^9 ly$ [34–36], the value of the fine structure constant α remains constant. More precisely, there are indications of a very slight variation of the parameter α.

The parameter α depends on the electron charge q, as given in

$$\alpha = \frac{q^2}{4\pi\varepsilon_0 c\hbar}. \tag{424}$$

Therefore, this parameter is not constant. We have

$$\frac{\alpha(r)}{\alpha} = \left(\frac{q(r)}{q}\right)^2.$$

Using equation (375) we also have

$$\frac{\alpha(r)}{\alpha} = \left(\frac{q(r)}{q}\right)^2 = \left(\frac{1-B}{1-B\exp\left(-\dfrac{k_1 r}{c}\right)}\right)^2. \tag{425}$$

From this equation it can be inferred that the parameter $\alpha(r)$ (essentially the electric charge$q(r)$), remains constant

for large distances r when the constant k_1 or the parameter B acquire extremely small values. According to relation (408) we have that

$$q(r \to \infty) \to q(1 - B).$$

This relation can be written as

$$\alpha(r \to \infty) \to \alpha(1 - B)^2.$$

Therefore, the value of the electric charge and of the parameter α in the very early Universe are only determined by the value of the parameter B. Hence, the parameter B has a very small value, independently of the value of constant k_1.

For very small values of the parameter B we see that

$$q(r \to \infty) \to q(1 - B) \to q.$$

This prediction does not cause any problems at the initial state of the Universe, since the electric charge exists in couples of opposite physical quantities. Such a relation cannot hold for the case of the rest mass, and indeed we know that

$$\frac{z}{1+z} < A < 1$$
$$m_0(r \to \infty) \to m_0(1 - A) \to 0.$$

From equation (376) we obtain $B > 0$. Thus, we arrive at the conclusion that the parameter$B$ acquires extremely small positive values.

The extremely small value of the parameter B assures the stability of the value of the parameter α for large distances r. Hence, we turn our attention not to the arithmetic value (which is likely to be extremely small, as is the case for the constant $k = \frac{1-A}{A}H$), but to the sign of the constant k_1.

For $k_1 > 0$ we obtain successively that

$$k_1 > 0$$
$$-\frac{k_1 r}{c} < 0$$
$$\exp\left(-\frac{k_1 r}{c}\right) < 1, \qquad (B > 0)$$
$$B\exp\left(-\frac{k_1 r}{c}\right) < B$$
$$-B\exp\left(-\frac{k_1 r}{c}\right) > -B$$
$$1 - B\exp\left(-\frac{k_1 r}{c}\right) > 1 - B, \qquad \left(1 - B\exp\left(-\frac{k_1 r}{c}\right) > 0\right)$$
$$\frac{1-B}{1 - B\exp\left(-\frac{k_1 r}{c}\right)} < 1.$$

From equation (425) we have that

$$\frac{\alpha(r)}{\alpha} = \left(\frac{q(r)}{q}\right)^2 < 1 \qquad k_1 > 0.$$

Therefore, for $k_1 > 0$ we will measure a slight decrease of the parameter α at large distances. Similarly, it turns out that for $k_1 < 0$ we will measure a slight increase of the parameter α at large distances [37]

$$\frac{\alpha(r)}{\alpha} = \left(\frac{q(r)}{q}\right)^2 > 1, \qquad k_1 < 0.$$

Based on the observational data we currently have, measurements of the variation of the parameter α have to be conducted for distances greater than 6×10^9 ly. The extremely small value of the (positive) parameter B renders these measurements difficult, in both cases.

7.11 The large structures in the Universe

The increase of the rest masses with the passage of time strengthens the gravitational interaction and accumulates matter towards various directions. The consequences of the accumulation of matter depend upon the quantity of the accumulated matter, as well as on the volume it occupies. In all cases, the total initial energy of the accumulated matter is zero, according to relation (407).

At large scales, at distances of order of magnitude 10^9 ly, the distribution of matter must have been determined by a large-scale destruction of the absolute homogeneity of the vacuum in the very early Universe. This explains the colossal webs of matter through vast expances of empty space that we observe with the modern observational instruments.

At smaller scales, within the dimensions of a galaxy, the accumulation of matter increases the temperature, as a result of the conversion of the gravitational potential energy into heat. A percentage of the particles of matter accumulates in a first central core of high temperature, while the remaining percentage remains distributed in the surrounding space during the period of accumulation. The slow rate at which the selfvariations occur, strengthens, also at a slow rate, the magnitude of the gravitational interaction, and allows a considerable percentage of the particles to remain in the surrounding space.

A further accumulation of the first core will lead to the formation of a second, more centralized core, until the temperature reaches the point where nuclear fusion starts. The initiation of nuclear fusion prevents the further accumulation of matter.

We separated the process of the accumulation into two phases, and we mentioned two cores for the following reason: The initial percentage of matter which remained outside the initial central core concerns the initial phase of the accumulation and is at a low temperature, slightly above $0K$ [38]. However, the percentage of matter which stays outside the second, and real central core, already has a high temperature. If we take into account the very high value of the Reynolds coefficient in this region, turbulent vortices will be generated. Therefore, the formation of stars should occur in this region.

In the final central core, the density of matter should be larger than in the rest of the galaxy. Clusters of galaxies are formed through similar processes.

Rough calculations give an equation correlating the mass and the volume of a galaxy [39–42]. This relation is consistent with the data we possess about galaxies (and galaxy clusters). But in reality, the process of accumulation is not separated into phases, but evolves in a continuous manner, from its beginning up to the formation of a galaxy. Therefore, we can only reach safe conclusions on the issue through computer simulations.

7.12 The origin of matter and the arrow of time

The equations of the theory of selfvariations predict at the limit, in the very distant past, that the beginning of the Universe was the vacuum. Therefore, we cannot consider a point to be the beginning of the universe, as proposed by the Standard Cosmological Model. All the points within the Universe are equivalent. The Universe originates "from everywhere", exactly as the cosmic microwave background radiation does (paragraph 7.8). Which physical mechanism can lead to such a result?

The theory of selfvariations predicts that the generalized particle can behave in such a way. The correlation of the vacuum with the condition $dS^2 = 0$ leads to such an interpretation, as we analyzed it in paragraph 5.9 and in paragraph 6.

What happens at the microcosm is a repetition at a local level, in a region of spacetime, of the condition that dominated throughout the spacetime occupied by the Universe during its emergence from the vacuum. That is how the slight perturbations of enormous spatial dimensions emerged within the initial homogeneity of the vacuum.

These perturbations were recorded on the cosmic microwave background radiation that followed (2.74 K) and which also originates from the whole Universe, as discussed in paragraph 7.8. Moreover, these perturbations are responsible for the large-scale distribution of matter in the Universe (paragraph 7.11).

The theory of selfvariations solves a fundamental problem of physical reality, which the physical theories of the last century are unable to solve. The equations of the theory of selfvariations include the arrow of time. The Universe originates from the vacuum and evolves towards a particular direction, which is determined by the selfvariations. The selfvariations continuously "distance" the Universe from the state of vacuum, but the Universe remains consistent with its origin:

The origin of matter from the vacuum, combined with the principles of conservation, has as a consequence that the energy content of the Universe is zero.

In the laboratory, the internality of the Universe to the process of measurement apparently "freezes" the time evolution of the selfvariations. On the contrary, the consequences of the selfvariations are directly imprinted on the observations we conduct at large distances. The Universe we observe today, and the complex processes taking place in Nature, are the results of the evolution of the selfvariations with the passage of time.

7.13 The future evolution of the Universe

The range of values parameter A takes is given by inequality (387)

$$\frac{z}{1+z} < A < 1.$$

Furthermore, equation (371) informs us that the parameter A approaches unity at an exceptionally slow rate, due to the extremely small value of the constant $k = \frac{1-A}{A}H$.

The parameter A appears in all of the equations we have stated. Because of this, the evolution of this parameter through time also determines the future evolution of the Universe, at least in the observations we will conduct in the far future.

From equation (388) we have that

$$\dot{H} = k\frac{\dot{A}(1-A) + A\dot{A}}{(1-A)^2}.$$

Using equation (371) we obtain

$$\dot{H} = k\frac{kA}{(1-A)^2}$$

$$\dot{H} = \frac{1}{A}\left(\frac{kA}{1-A}\right)^2$$

$$\dot{H} = \frac{1}{A}H^2.$$

For $A \sim 1, H = 60\frac{\text{km}}{\text{sMpc}} = 2\times 10^{-18}\ \text{s}^{-1}$

$$\dot{H} = 4\times 10^{-36}\text{s}^{-2}.$$

The Hubble parameter varies at an extremely slow rate.

We shall now see how the redshift z varies with the passage of time. From equation (382) we get

$$z = \frac{1 - A\exp\left(-\frac{kr}{c}\right)}{1-A} - 1$$

$$z = \frac{A}{1-A}\left(1 - \exp\left(-\frac{kr}{c}\right)\right). \tag{426}$$

For the same distance r we have that

$$\dot{z} = \left(\frac{A}{1-A}\right)^{\bullet}\left(1 - \exp\left(-\frac{kr}{c}\right)\right)$$

$$\dot{z} = \frac{\dot{A}}{(1-A)^2}\left(1 - \exp\left(-\frac{kr}{c}\right)\right).$$

Using equation (371) we see that

$$\dot{z} = \frac{kA}{(1-A)^2}\left(1 - \exp\left(-\frac{kr}{c}\right)\right).$$

Considering equation (426) we obtain

$$\dot{z} = \frac{k}{1-A}z$$
$$\dot{z} = \frac{1}{A}\frac{kA}{1-A}z.$$

Through equation (388) we arrive at

$$\dot{z} = \frac{H}{A}z. \tag{427}$$

For $H = 2 \times 10^{-18}\mathrm{s}^{-1} = 6.3 \times 10^{-11}\mathrm{year}^{-1}$ and $A \sim 1$ we obtain

$$\dot{z} = z \cdot 6.3 \times 10^{-11}\mathrm{year}^{-1}. \tag{428}$$

The rate of increase of the redshift z is a measure with which to evaluate the future evolution of the Universe.

8 The Topographic Theorem

For a material point particle, the velocity υ of the selfvariations is defined by equation (6)

$$\upsilon = \frac{c}{r}\boldsymbol{r}. \tag{429}$$

This equation refers solely to the material point particle. On the contrary, equation (33)

$$\frac{\boldsymbol{\upsilon}}{c} = \begin{bmatrix} \cos\delta \\ \sin\delta\cos\omega \\ \sin\delta\sin\omega \end{bmatrix} \tag{430}$$

has more general validity. The velocity in equation (430) satisfies the relation $\|\upsilon\| = c$, without necessarily having the form (429). Therefore, we have to study the properties of the velocity υ, as they follow from equation (430). The differentiation between the two equations occurs in equations (43) and (44)

$$\nabla\delta = \lambda_1\frac{\boldsymbol{\upsilon}}{c} + K\boldsymbol{\beta} + L\boldsymbol{\gamma}$$
$$\nabla\omega = \lambda_2\frac{\boldsymbol{\upsilon}}{c} + M\boldsymbol{\beta} + N\boldsymbol{\gamma}$$

which take the form

$$\nabla\delta = -\frac{\partial\delta}{c\partial t}\frac{\boldsymbol{\upsilon}}{c} + K\boldsymbol{\beta} + L\boldsymbol{\gamma}$$
$$\nabla\omega = -\frac{\partial\omega}{c\partial t}\frac{\boldsymbol{\upsilon}}{c} + M\boldsymbol{\beta} + N\boldsymbol{\gamma}. \tag{431}$$

We will mention the general properties of the velocity υ, without citing the relevant proofs.

The coefficients $\frac{\partial\delta}{c\partial t}$, K, L, $\frac{\partial\omega}{c\partial t}$, M, N are not independent from each other, but are constrained by the following compatibility equations:

$$\begin{aligned}
&\frac{\partial\delta}{c\partial t}(L - M\sin\delta) + (KM + LN)\cos\delta - \boldsymbol{\gamma}\cdot\nabla K + \\
&+\boldsymbol{\beta}\cdot\nabla L = 0 \\
&\frac{\partial\omega}{c\partial t}(L - M\sin\delta) + \left(M^2 + N^2\right)\cos\delta - \boldsymbol{\gamma}\cdot\nabla M + \\
&+\beta\cdot\nabla N = 0 \\
&\frac{\partial K}{\partial t} + \boldsymbol{\upsilon}\cdot\nabla K = -c\left(K^2 + LM\sin\delta\right) \\
&\frac{\partial L}{\partial t} + \boldsymbol{\upsilon}\cdot\nabla L = -cL(K + N\sin\delta) \\
&\frac{\partial M}{\partial t} + \boldsymbol{\upsilon}\cdot\nabla M = -cM(K + N\sin\delta) \\
&\frac{\partial N}{\partial t} + \boldsymbol{\upsilon}\cdot\nabla N = -c\left(LM + N^2\sin\delta\right) \\
&\frac{\partial}{\partial t}\left(\frac{\partial\delta}{c\partial t}\right) + \boldsymbol{\upsilon}\cdot\nabla\left(\frac{\partial\delta}{c\partial t}\right) = -K\frac{\partial\delta}{\partial t} - L\sin\delta\frac{\partial\omega}{\partial t} \\
&\frac{\partial}{\partial t}\left(\frac{\partial\omega}{c\partial t}\right) + \boldsymbol{\upsilon}\cdot\nabla\left(\frac{\partial\omega}{c\partial t}\right) = -M\frac{\partial\delta}{\partial t} - N\sin\delta\frac{\partial\omega}{\partial t}.
\end{aligned} \tag{432}$$

These equations are valid in every inertial frame of reference.

For the inertial reference frames S and S', as we defined them in paragraph 3, the following Lorentz-Einstein transformations hold

$$\begin{aligned}
&\frac{\partial\delta'}{c\partial t'} = \frac{\partial\delta}{c\partial t} + \frac{u}{c}\frac{K\sin\delta}{1 - \frac{u}{c}\cos\delta} \\
&K' = \frac{K}{\gamma\left(1 - \frac{u}{c}\cos\delta\right)} \qquad L' = \frac{L}{\gamma\left(1 - \frac{u}{c}\cos\delta\right)} \\
&\frac{\partial\omega'}{c\partial t'}\sin\delta' = \frac{\partial\omega}{c\partial t}\sin\delta - \frac{u}{c}M\sin\delta\frac{\sin\delta}{1 - \frac{u}{c}\cos\delta} \\
&M'\sin\delta' = \frac{M\sin\delta}{\gamma\left(1 - \frac{u}{c}\cos\delta\right)} \\
&N'\sin\delta' = \frac{N\sin\delta}{\gamma\left(1 - \frac{u}{c}\cos\delta\right)}.
\end{aligned} \tag{433}$$

We define the vector

$$\begin{aligned}
\boldsymbol{t} &= \nabla\delta\times\sin\delta\nabla\omega = t_1\tfrac{\upsilon}{c} + t_2\boldsymbol{\beta} + t_3\boldsymbol{\gamma} = \\
&= (KN\sin\delta - LM\sin\delta)\frac{\boldsymbol{\upsilon}}{c} + \\
&+\left(\frac{\partial\delta}{c\partial\boldsymbol{t}}N\sin\delta - L\frac{\partial\omega}{c\partial t}\sin\delta\right)\boldsymbol{\beta} \\
&+\left(K\frac{\partial\omega}{c\partial t}\sin\delta - \frac{\partial\delta}{c\partial t}M\sin\delta\right)\boldsymbol{\gamma}.
\end{aligned} \tag{434}$$

The topography of the generalized photon is defined by the following theorem:

Theorem 5. *The Topographic Theorem. For every inertial frame of reference and for every generalized photon, the following hold:*

1. *If it is* $(t_1, t_2, t_3) \neq (0, 0, 0)$*, then the generalized photon is of one spatial dimension. The material points of the generalized photon are arranged on a curve. At each point of the curve the vector t is tangent on the curve.*
2. *The generalized photon can have two spatial dimensions, with its material points arranged on a surface. Then at each point of the surface, the vector n, vertical to the surface, is given by* $\boldsymbol{n} = \frac{\nabla\delta}{\|\nabla\delta\|} = \frac{\nabla\omega}{\|\nabla\omega\|}$.
3. *If the material points of the generalized photon are arranged in the three-dimensional space, then it is* $K = L = M \sin\delta = N \sin\delta = \frac{\partial\delta}{c\partial t} = \frac{\partial\omega}{c\partial t} \sin\delta = 0$.

For the material point particle and for the velocity vector (429), we obtain from equations (51) and (34)

$$\frac{\partial\delta}{c\partial t} = -\frac{\boldsymbol{u}\cdot\boldsymbol{\beta}}{cr\left(1-\dfrac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}\right)}$$

$$K = \frac{1}{r} \qquad L = 0$$

$$\sin\delta\frac{\partial\omega}{c\partial t} = -\frac{\boldsymbol{u}\cdot\boldsymbol{\gamma}}{cr\left(1-\dfrac{\boldsymbol{v}\cdot\boldsymbol{u}}{c^2}\right)}$$

$$M\sin\delta = 0 \qquad N\sin\delta = \frac{1}{r}.$$

Thus, we get

$$t_1 = KN\sin\delta - LM\sin\delta = \frac{1}{r^2} \neq 0$$

and, therefore, it is $t \neq (0, 0, 0)$. Consequently, in the case of equation (429) the generalized photon is of one spatial dimension. Therefore, the trajectory representation theorem emerges, as we saw in paragraph 2.4.

The topographic theorem permits the study of the self-variations for non-point material particles.

Submitted on: March 03, 2013 / Accepted on March 08, 2013

References

1. Mase C.C. Continuum Mechanics. 1970, Schaum Publ. Co.
2. Hunter S.C. Mechanics of Continuous Media. 1976, Halsted Press.
3. Wallace P.R. Mathematical Analysis of Physical Problems. 1984, New York, Dover.
4. Spivak M. A Comprehensive Introduction to Differential Geometry. 1999.
5. French A.P. Special Relativity. 1968, Nelson, London.
6. Bohm D. The special Theory of Realitivity. 1965, Benjamin, New York.
7. Couderc P. La Relativite. 1963, Presses Univeritaires de France.
8. Good R.H. Basic concepts of Relativity. 1968, Reinhold book corporation.
9. Rindler W. Relativity. Special, General and Cosmological. 2001, Oxford University Press.
10. Rindler W. Introduction to Special Relativity. 1991, Clarendon Press, Oxford.
11. Synge J.L. Relativity: The Special Theory. 1972, North-Holland Publishing Company.
12. Landau L.D., Lifshitz E.M. Fluid Mechanics, Vol. 6, 1987, Butterworth-Heinemann.
13. Wang C.C. Mathematical Principles of Mechanics and Electromagnetism. Part A: Analytical and Continuum Mechanics, 1979, Plenum Press.
14. Schwartz M. Principles of Electrodynamics. 1987, New York, Dover.
15. Shadowitz A. The Electromagnetic Field. 1988, New York, Dover.
16. Griffiths D.J. Introduction to Electrodynamics. 1989, Englewood Cliffs, NJ: Prentice Hall, p. 368.
17. Reitz J.R., Milford F.J., Christy R.W. Foundations of Electromagnetic Theory. 2000, MA: Addison-Wesley.
18. Landau L.D., Lifshitz E.M., Pitaevskii L.P. Electrodynamics of Continuous Media. 1984, Butterworth-Heinemann.
19. Flugge W. Tensor Analysis and Continuum Mechanics. 1972, Springer-Verlag.
20. Davies W.R. Classical Fields, Particles and the Theory of Relativity. 1970, Gordon and Breach.
21. Corson E.M. Introduction to Tensors, Spinors, and Relativistic Wave-Equations. 1953, Blackie and Son.
22. French A. An introduction to Quantum Mechanics. 1978, W.W. Norton.
23. Feynman R. Lectures on Physics, Vol. III, 1965, Addison-Wesley.
24. Bohm D. Quantum Theory, 1951, Prentice-Hall.
25. Dicke R.H., Wittke J.P. Introduction to Quantum Mechanics. 1960, Addison-Wesley.
26. Gasiorowitz S. Quantum Physics. 1974, Wiley International.
27. Dirac P.A.M. The Principles of Quantum Mechanics. 1967, Oxford University Press.
28. Riess A.G. et al. Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant, *Astronomical Journal*, 1998, v. 116, 1009–1038.
29. Perlmutter S., Turner M.S., White M. Constraining Dark Energy with Type Ia Supernovae and Large-Scale Structure. *Physical Review Letters*, 1999, v. 83, 670–673.
30. Baldi M., Pettorino V. High-z massive clusters as a test for dynamical coupled dark energy. *MNRAS*, 2011, v. 412:L.
31. Rybicki G.B., Lightman A.P. Radiative Processes in Astrophysics. 1985, John Wiley-Interscience.
32. Malcolm S.L. High Energy Astrophysics. 2011, University of Cambridge.
33. Spergel D.N. et al. Wilkinson Microwave Anisotropy Probe (WMAP) Three Year Results: Implications for Cosmology. 2007, ApJS 170, 377.
34. Clifford M.W. Theory and Experiment in Gravitational Physics. 1981, Cambridge University Press.
35. Webb J.K., Flambaum V.V., Churchill C.W., Drinkwater M.J., Barrow J.D. Search for time variation of the fine structure constant. *Physical Review Letters*, 1999, v. 82, 884–887.
36. Tzanavaris P., Webb J.K., Murphy M.T., Flambaum V.V., Curran S.J. Limits on variations in fundamental constants from 21-cm and ultraviolet Quasar absorption lines. *Physical Review Letters*, 2005, v. 95, pp. 41301-41304.

37. Webb J.K., King J. A., Murphy M. T., Flambaum V. V., Carswell R. F., Bainbridge M. B., Indications of a Spatial Variation of the Fine Structure Constant. *Physical Review Letters*, 2011, v. 107, pp. 191101-1-191101-5.

38. Knebe A., Devriendt J.E.G., Gibson B.K., Silk J. Top-down fragmentation of warm dark matter filament. *Monthly Notices of the Royal Astronomical Society*, 2003, v. 345, 1285.

39. Manousos E. The theory of self-variations. A continuous slight increase of the charges and the rest masses of the particles can explain the cosmological data. *Nuovo Cimento B*, 2007, v. 122, 359-388.

40. Gott J.R. et al. A Map of the Universe. 2005, ApJ, 624, 463.

41. Labini F.S., Vasilyev N.L., Baryshev Y.V. Persistent fluctuations in the distribution of galaxies from the Two-degree Field Galaxy Redshift Survey. *EPL*, 2009, v. 85, 29002.

42. Peacock J.A. et al. A measurement of the cosmological mass density from clustering in the 2dF Galaxy Redshift Survey. *Nature*, 2001, v. 410, 169–173.

On the Uniform Dimension System. Is There the Necessity for Coulomb?

Anatoly V. Belyakov

E-mail: belyakov.lih@gmail.com

The dimensions of electrical units (Ampere, Coulomb, etc.) are surplus. It is shown that the most appropriate is to replace the electric charge with the ultimate momentum of the electron. Then all the dimensions of electrical and magnetic values get simplified and assume physically obvious form.

Although electric and magnetic dimensions in systems CGSE and CGSM are expressed in the terms of mass, length, and time units (in SI units Ampere was added), they seem strange and bizarre. The exception is a unit of capacity in the system CGSE whose dimension (centimeters) looks convincing. Of course, the dimensions are relative; however, it causes internal resistance, misunderstanding, and difficulties in the perception of the relevant areas of physics, especially for students. Is there need of having electrical values proper, above all Ampere (or Coulomb)?

Indeed, the basic formulae (the electrical force between the charges and the magnetic force between current-carrying conductors) can be represented with a single dimensional factor of force. Only the number of electric charges z is meaningful for the force of electrical and magnetic interaction. In the Coulomb formula a unit of the charge can be expressed through the electron mass m_e and the classical electron radius r_e. Then Coulomb formula can be obtained as:

$$F_e = \frac{m_e c^2}{r_e}\left(\frac{r_e}{r}\right)^2 z_1 z_2 ,$$

where r, c, z_1, z_2 are, respectively, the distance between the charges, the velocity of light, and the number of the electric charges.

Here dimensional coefficient $m_e c^2/r_e$ is the centrifugal force that occurs when an electron moves with the light velocity c of the radius r_e. This force is equivalent to the force acting between two elementary charges by the given distance and its numerical value is a very ordinary magnitude equal 29.06 N.

In what units electric charge should be measured? According to John A. Wheeler' idea, the charged microparticles are special points in the three-dimensional spatial surface of our world, connected to each other through "wormholes" — vortical tubes analogous to the lines of current working according to the "input-output" ("source-drain") principle, but in an additional dimension of space (but that does not mean that it is necessary to add a fourth spatial dimension [1]). In this model the electric charge is not a special kind of matter: the electric charge only manifests the degree of the non-equilibrium state of physical vacuum; it is proportional to the momentum of physical vacuum in its motion along the contour of the vortical current tube.

Therefore, the most appropriate is to replace the electric charge in formulae Coulomb and Ampere with the ultimate momentum of the electron $m_e c$. Then all the dimensions of electrical and magnetic values get strikingly simplified and assume sensible and physically obvious form. So, in SI units: current becomes force — [kgm/sec^2] or [N], the potential becomes velocity — [m/sec], capacity becomes mass of the electrons accumulated on the plates of the capacitor — [kg], conductivity becomes mass velocity — [kg/sec], inductance becomes the reciprocal value of mass acceleration — [sec^2/kg], the magnitude of the solenoid magnetic field becomes the number of turns per unit of solenoid length — [m^{-1}], etc.

The numerical values of the expressions for the electrical and magnetic forces, written in a "Coulombless" form with the charge replaced by the ultimate momentum of the electron, coincide with these values based on standard expressions at the following conditions:

— the value $4\pi\varepsilon_0$, which in SI units is 1.11×10^{-10} Farad/metre, is replaced by a new electric constant $\varepsilon_0 = m_e/r_e = 3.23\times10^{-16}$ [kg/m];

— magnetic constant μ_0, which in SI units is $4\pi\times10^{-7}$ Henry/metre, is replaced by a new magnetic constant $\mu_0 = 1/\varepsilon_0 c^2 = 0.0344$ [N^{-1}].

Thus, the electric constant ε_0 makes sense the linear density of the vortex tube current, and the magnetic constant μ_0 makes sense the reciprocal value of the interaction force between two elementary charges.

With such mechanistic interpretation Wheeler's scheme numerical values of the electric charge and radiation constants were successfully obtained [2]. In such system value $587\, m_e c$ [kgm/sec] is the momentum of the vortical tube current the whole, it numerically corresponds to the electron charge $e_0 = 1.602\times10^{-19}$ Coulomb; at the same time the value $m_e c$ [kgm/sec] corresponds to the "point-like" electron charge. The value of 587 [m/sec] corresponds to one Volt in SI units, the value of $4\pi/587^2 = 3.6\times10^{-5}$ [kg] corresponds to one Farad in SI units etc. Thus there would be no need for the systems of CGSE units, CGSM units and the Gaussian units. Replacing the dimensions and introducing new electromagnetic constants is a purely technical problem, although it is hardly practicable today. It is more important that mechanistic interpretation of the electromagnetic parameters

reveals theirs physical meaning and gives help understanding the nature of phenomena.

For example, the electrical capacity of a sphere of radius $R = 1$ [m] is equal to $3.23{\times}10^{-16}$ [kg], which corresponds to $3.5{\times}10^{14}$ electrons distributed over its surface and be responsible for the movement of the charges, while the average distance between the charges is $r = \pi\sqrt{Rr_e} = 1.67{\times}10^{-7}$ [m].

The capacitor charge is proportional to the potential, and it is easy to determine the satiation potential φ when the number of electric charges becomes equal with the number of electrons carrying the charges. Since the "point-like" electron cannot carries momentum over $m_e c$, for an only charge: $\varphi = m_e c/m_e = c$ [m/sec] or, in SI units, $c/587 = 511000$ Volts. If this potential is exceeded the mentioned magnitude, the charge is spontaneously flowing into the surrounding space.

In these examples (and in others) quite reasonable values have been obtained, which could not be if all above-stated would be wrong. Of course, such associations of electrical and magnetic values with mechanical ones do not yet mean reduction of electromagnetic phenomena to mechanical ones. The question immediately arises, how does the electron be able to carry momentum which exceeds many times your own one? However, this question implicitly always existed because the term "charge" is, in fact, the delicate symbol of not properly understood electricity essence. To some extent, the response has been received in this article, as well as in [2, 3].

Submitted on June 05, 2013 / Accepted on June 16, 2013

References

1. Belyakov A.V. On the Independent Determination of the Ultimate Density of Physical Vacuum. *Progress in Physics*, 2011, v.2, 27–29.
2. Belyakov A.V. Charge of the electron, and the constants of radiation according to J. A. Wheeler's geometrodynamic model. *Progress in Physics*, 2010, v.4, 90–94.
3. Belyakov A.V. Macro-analogies and gravitation in the micro-world: further elaboration of Wheeler's model of geometrodynamics. *Progress in Physics*, 2012, v.2, 47–57.

Further Problems with Integral Spin Charged Particles

E. Comay

Charactell Ltd., PO Box 39019, Tel-Aviv, 61390, Israel. E-mail: elicomay@post.tau.ac.il

The structure of the Lagrangian density of quantum theories of electrically charged particles is analyzed. It is pointed out that a well known and self-consistent expression exists for the electromagnetic interactions of a spin-1/2 Dirac particle. On the other hand, using the Noether theorem, it is shown that no such expression exists for the spin-0 Klein-Gordon charged particle as well as for the $W^{\pm}$ spin-1 particle. It is also explained why effective expressions used in practical analysis of collider data cannot be a part of a self-consistent theory. The results cast doubt on the validity of the electroweak theory.

1 Introduction

Since its very beginning, quantum theory has provided expressions describing electromagnetic interactions. In particular, the Dirac equation of spin-1/2 charged particle takes a covariant form [1, see pp. 16–24]. As is well known, electromagnetic interactions of a Dirac particle have an extraordinary experimental support. Later, a quantum theory of a spin-0 Klein-Gordon (KG) charged particle was published [2, see pp. 188–205]. In the electroweak theory which was constructed several decades later, the $W^{\pm}$ spin-1 charged boson plays a cardinal role. The discussion presented in this work examines the Lagrangian density of quantum theories. As is well known, the electromagnetic interaction term of these theories depends on a contraction of the charged particle's 4-current and the external 4-potential $j_\mu A^\mu$. Thus, the Noether theorem is used for deriving expressions for the charged particle's 4-current. In this way the analysis proves that electromagnetic theories of spin-0 and spin-1 particles contain inherent contradictions.

Units where $\hbar = c = 1$ are used in this work. Hence, only one dimension is required and it is the length, denoted by [L]. For example, mass, energy and momentum have the dimension $[\mathrm{L}^{-1}]$, etc. Greek indices run from 0 to 3 and the diagonal metric used is $g_{\mu\nu} = (1, -1, -1, -1)$. The symbol $_{,\nu}$ denotes the partial differentiation with respect to x^ν. The summation convention is used for Greek indices. The second section presents theoretical elements that are used in the discussion. The third section contains a proof showing that electromagnetic interactions cannot be a part of a self-consistent theory of spin-0 and of spin-1 quantum particles. Concluding remarks can be found in the last section.

2 The theoretical basis of the analysis

The following discussion examines the structure of a quantum theory of an electrically charged particle and its interaction with electromagnetic fields. The need for a Lagrangian density as basis for a relativistic quantum theory has become a common practice. This issue can be derived from the fact that the phase is an argument of an exponent. Thus, the power series expansion of the argument proves that the phase must be a dimensionless Lorentz scalar. This requirement is satisfied if the action (divided by $\hbar$) is used for the phase and the Lagrangian density is a Lorentz scalar whose dimension is $[\mathrm{L}^{-4}]$. Indeed, in this case, the action

$$S = \int \mathcal{L}\, d^4x \tag{1}$$

is a dimensionless Lorentz scalar.

The form of the required Lagrangian density is

$$\mathcal{L}(\Phi^\dagger, \Phi^\dagger_{,\mu}, \Phi, \Phi_{,\mu}, A^\mu, F^{\mu\nu}), \tag{2}$$

where Φ denotes the function of the charged quantum particle and $A^\mu, F^{\mu\nu}$ denote the electromagnetic 4-potential and its fields, respectively. In the discussion presented herein the quantum function Φ represents either scalar, spinor or vector particle. In specific cases the notation ϕ represents a KG charged particle, ψ denotes a Dirac particle and W^μ denotes the $W^{\pm}$ particles. Evidently, (1) and (2) prove that the function Φ has dimension.

Maxwellian electrodynamics is derived from the following Lagrangian density [3, see pp. 71–81]

$$\mathcal{L} = -\frac{1}{16\pi} F_{\mu\nu}F^{\mu\nu} - j_\mu A^\mu, \tag{3}$$

where j^μ denotes the charge's 4-current and the last term of (3) represents the electromagnetic interaction.

This expression demonstrates the crucial role of the 4-current in a self-consistent theory of an electrically charged particle. As is well known, the charge 4-current must satisfy the continuity equation

$$j^\mu_{,\mu} = 0. \tag{4}$$

The standard method used for constructing such a 4-current relies on Noether's theorem [4, see p. 20]. Thus, in the present case, the expression for the 4-current boils down to the following form

$$j^\mu = i\frac{\partial \mathcal{L}}{\partial \Phi^\dagger_{,\mu}}\Phi^\dagger - i\frac{\partial \mathcal{L}}{\partial \Phi_{,\mu}}\Phi. \tag{5}$$

(Note that due to the opposite phase sign of $\Phi^\dagger$ and Φ, corresponding terms derived from these functions have opposite

sign.) Thus, in the case of a charged particle, the Noether 4-current (5) is multiplied by the electric charge e. Relying on (5), one concludes that the 4-current is derived from terms of the Lagrangian density that contain a derivative of the field function with respect to the coordinates x^μ. The 0-component of (5) represents the particle's density. Hence, the dimension of j^μ is $[\mathrm{L}^{-3}]$.

A standard method used for the introduction of electromagnetic interaction is to substitute the following transformation in the free Lagrangian density of the particle (see e.g. [1, p. 10])

$$-i\frac{\partial}{\partial x^\mu} \to -i\frac{\partial}{\partial x^\mu} - eA_\mu(x^\nu). \tag{6}$$

Later, this substitution is called the standard form of electromagnetic interaction. This form as well as other forms of electromagnetic interactions are discussed in the next section.

3 Quantum charged particles

The Dirac Lagrangian density of a free spin-1/2 particle is [4, see p. 54]

$$\mathcal{L} = \bar{\psi}\left[\gamma^\mu i\partial_\mu - m\right]\psi. \tag{7}$$

This expression is linear in the mass. Hence, the dimension $[\mathrm{L}^{-4}]$ of the Lagrangian density means that the dimension of the Dirac function ψ is $[\mathrm{L}^{-3/2}]$. An application of the Noether relation (5) for a construction of the 4-current yields the well known Dirac expression [1, see pp. 23–24] which is written below in the standard notation

$$j^\mu = e\,\bar{\psi}\gamma^\mu\psi. \tag{8}$$

The dimension $[\mathrm{L}^{-3/2}]$ of the Dirac function ψ shows that (8) has the required dimension.

The case of the KG and of the W Lagrangian density is different. Here the mass term takes the form (see [4, p. 26] and [5, p. 309], respectively)

$$-m^2\Phi^\dagger\Phi. \tag{9}$$

Different numerical factors of (9) are not mentioned and the same is true for the contraction of the 4 components of the W function. Relationship (9) means that the dimension of the KG and of the W functions is $[\mathrm{L}^{-1}]$. Thus, in order to satisfy the $[\mathrm{L}^{-4}]$ dimension of the Lagrangian density of these particle, it must contain terms that are *bilinear* in derivatives with respect to the space-time coordinates x^μ. Applying the Noether relation for the 4-current (5), one finds that *the 4-current of the KG and of the W particles contains a derivative with respect to* x^μ. This property means that utilizing of the standard form of the introduction of electromagnetic interactions (6), one finds that *the 4-current of the KG and of the W particles depends linearly on the 4-potential of the electromagnetic fields.* (This is certainly inconsistent with gauge invariance, because here a gauge transformation alters charge density and the associated field values as well. However, this matter is not discussed in the present work.) The dependence of the charged KG 4-current on the external electromagnetic 4-potential has already been shown a long time ago [2, see p. 199].

Let us turn to the electromagnetic fields. The interaction term of the Maxwellian Lagrangian density (3) is $j_\mu A^\mu$. Now, if the 4-current j^μ of the KG and of the W particles depends linearly on the 4-potential of electromagnetic fields then *there is a quadratic term of the 4-potential in the expression for the interaction term in the Maxwellian Lagrangian density (3).* This is a contradiction because in Maxwellian electrodynamics the interaction term must be linear in the 4-potential [3, see pp. 78–79].

The foregoing discussion proves that there is no theoretically valid expression for the electromagnetic interaction of a KG particle and of the W boson as well. Thus, in the case of the W boson people resort to a phenomenological expression that goes by the name *effective Lagrangian density* [6, 7]. Using standard notation for the W field, one of the nonvanishing electromagnetic interaction terms of the effective Lagrangian density is

$$\mathcal{L}_{int} = -ie\left(W^\dagger_{\mu\nu}W^\mu A^\nu - W^\dagger_\mu W^{\mu\nu}A_\nu\right). \tag{10}$$

The articles [6,7] have been cited many times and (10) is still used in a collider data analysis [8, see eq. (1)] [9, see eq. (3)].

The following argument proves that (10) is indeed an effective expression which cannot be justified theoretically. Let us assume that (10) is a term in a theoretically justifiable Lagrangian density. In this case the following expression

$$j^\nu = -ie\left(W^{\dagger\nu}_\mu W^\mu - W^\dagger_\mu W^{\mu\nu}\right) \tag{11}$$

represents the electric 4-current of the W boson. But (11) contains the factors $W^{\dagger\mu\nu}$ and $W^{\mu\nu}$, and by the definition $W_{\mu\nu} = \partial_\mu W_\nu - \partial_\nu W_\mu$, each of which is a derivative with respect to x^μ. Therefore, due to the Noether theorem (5), the interaction term (10) alters the 4-current of the W boson and adds to it a troublesome term that is proportional the the external electromagnetic 4-potential A^μ. Hence, *contrary to the assumption examined herein, (11) does not represent the 4-current of the W boson.* This contradiction substantiates the proof.

A second electromagnetic term which is introduced into the effective Lagrangian density of the W is [6–9]

$$\mathcal{L}_{int} = ie\,W^\dagger_\mu W_\nu F^{\mu\nu}. \tag{12}$$

This term is certainly inconsistent with electromagnetic interactions because these interactions are proportional to the 4-current of the charged particle and the dimension of the 4-current is $[\mathrm{L}^{-3}]$. On the other hand, it is proved above that the dimension of the W function is $[\mathrm{L}^{-1}]$ and that of $W^\dagger_\mu W_\nu$ is $[\mathrm{L}^{-2}]$. Therefore, (12) cannot represent a consistent electromagnetic interaction.

4 Conclusions

The solid mathematical structure of the spin-1/2 Dirac equation and its successful experimental status are pointed out above. Here a self-consistent relativistically covariant electromagnetic interaction exists. Thus, nobody finds the need to resort to "effective Lagrangian density".

A different situation holds for the cases of spin-0 and spin-1 elementary particles. It is proved in this work that for these particles the standard methods used for constructing electromagnetic interactions fail. Furthermore, it is proved above that the authors of [6, 7] are right in their description of the W boson electromagnetic interaction (10) as an effective expression. However, a proof presented in the previous section shows that (10) cannot be a part of a theoretically self-consistent Lagrangian density. This outcome means that *the W boson cannot carry an electric charge.* Now, the W boson takes a vital part in the unification of electrodynamics with weak interaction which is called electroweak theory. Therefore, the results cast doubt on the validity of the electroweak theory.

Another result of the discussion presented above is that the experimentally detected W boson cannot be an elementary particle described by a field function that takes the form $W^{\pm\mu}(x^{\nu})$. Indeed, a dependence on a single set of space-time coordinates x^{μ} is a property of a structureless pointlike elementary particle like the electron etc. Thus, the actual $W^{\pm}$ particles must be composite particles and it looks plausible to regard them as a combination of mesons of the top quark and either of the d, s, b antiquarks or vice versa. It turns out that the conclusions of this work provide an independent support to similar conclusions that have been published earlier [10]. It should also be noted that the results of this work are consistent with Dirac's lifelong objection to the KG equation [11].

Submitted on May 23, 2013 / Accepted on June 10, 2013

References

1. Bjorken J. D. and Drell S. D. Relativistic Quantum Mechanics. McGraw-Hill, New York, 1964.
2. Pauli W. and Weisskopf V. The quantization of the scalar relativistic wave equation. *Helvetica Physica Acta*, 1934, v. 7, 709–731. English translation in Miller A. I. Early Quantum Electrodynamics. University Press, Cambridge, 1994. (In the text, page references apply to the English translation.)
3. Landau L. D. and Lifshitz E. M. The Classical Theory of Fields. Elsevier, Amsterdam, 2005.
4. Bjorken J. D. and Drell S. D. Relativistic Quantum Mechanics. McGraw-Hill, New York, 1965.
5. Weinberg S. The Quantum Theory of Fields, Vol II. Cambridge University Press, Cambridge, 1996.
6. Hagiwara K., Peccei R. D., Zeppenfeld D. and Hikaso K. *Nuclear Physics*, 1987, v. B282, 253.
7. Hagiwara K., Woodside J., Zeppenfeld D. *Physical Review*, 1990, v. D41, 2113.
8. Abazov V. M. et al (D0 collaboration). *Physics Letters*, 2012, v. B718, 451.
9. Aad G. et al (ATLAS collaboration). *Physics Letters*, 2012, v. B712, 289.
10. Comay E. Quantum Constraints on a Charged Particle Structure. *Progress in Physics*, 2012, v. 4, 9–13.
11. Dirac P. A. M. Mathematical Foundations of Quantum Theory. In Marlow A. R. (ed.) Mathematical Foundations of Quantum Theory. Academic, New York, 1978.

Studies of Pulsed Signals in High-precision Experiments (Antarctica)

Sergey N. Shapovalov*, Oleg A. Troshichev†, Vacheslav I. Povazhny‡ and Igor V. Moskvin**

SSC Arctic and Antarctic Research Institute. 38 Beringa st., 199397 St-Petersburg, Russia
*E-mail: shapovalov@aari.ru; †E-mail: olegtro@aari.ru; ‡E-mail: povazviach@gmail.com; **E-mail: im_geo@aari.ru

The paper presents the results of studies on pulsed signals in photocurrent (PCC-2 instrument), in the 565-nm LED spectrum, and in the atmospheric zenith spectrum (342.5 nm). According to the results of statistical analysis of data measurements for the period from 24.04.04 till 01.02.06 a correlation between the temporal distribution of pulsed signals in photocurrent PCC-2 and CA F10.7 cm (2800 MHz) index and the total solar radiation (TSI) was established. In the course of the parallel measurements of photocurrent in PCC-2 and fluctuations in the spectra frequencies of the LED and the atmosphere zenith, based on the average daily values of the standard deviation, the identical trend in the photocurrent pulse signals (PCC-2) and the fluctuations at 520-nm LED spectrum and 342.5-nm atmosphere zenith spectrum was detected (AvaSpec–2048 spectrometer).

1 Introduction

The way towards recognition of the role of unknown cosmophysical effects on the Earth processes presents certain difficulties. At the first stage of research, the existence of non-electromagnetic radiation affecting the physical and biological systems was hypothesized. Among these, the conclusion about the advanced (4–6 days lead time) increase of corynebacteria sensitivity to the emergence of active formations on the surface of the Sun made by A. L. Chizhevsky and S. T. Velkhover [1] should be mentioned. It may be only assumeday be tions ioned that these formations are linked to perturbations in the deep spheres of the Sun and are accompanied by the topography changes in its gravitational field. At least, this is supported by the existence of lead time phenomenon undetected by other methods.

The applied research of the cosmophysical radiation and its impact on physical systems started with the works of N. A. Kozyrev [2], who had registered with a telescopic system the effect of unknown factor of high penetrating power. Due to the fact that the optical entrance of the telescope was overlaid with a metal screen, a non-electromagnetic origin of the registered radiation may be suggested. The results received by N. A. Kozyrev were confirmed later, in the experiments of the workgroup headed by M. M. Lavrentiev, Fellow Russian Academy [3]. Valuable results were obtained at the recent stage of research [4–8].

2 Studies of pulsed signals in photocurrent measurements with PCC-2

Technical characteristics of PCC-2 instrument: Photoelectric concentration colorimeter (PCC-2) is designed for measuring coefficients of transmission and optical density of solutions in the range of 315–980 nm (a set of optical light filters), as well as to determine the concentration of substances in solution by constructing calibration curves. Radiation detectors: F-26 photocell for operating in 315–540 nm range and FD-7K photodiode for operating in 590–980 nm range (Fig. 1).

Fig. 1: Photoelectric concentration colorimeter (PCC-2).

Recording device of the instrument is M 907-10 microammeter with digitized scale for coefficients of transmission and optical density. Power supply 220 ± 22 V, $50/60 \pm 0.5$ Hz. The source of radiation – KGM $6.3-15$ small-size halogen lamp. The range of readings that characterizes random errors does not exceed 0.3%.

The normal running conditions for PCC-2 are: temperature $(20 \pm 5)^{\circ}$C, relative humidity 45–80%, mains voltage 220 ± 4.4V, 50 Hz.

In the course of the Antarctic expedition to Mirny station, 1996–1997, during measurements of the dynamics rate of biochemical reactions [9], sharp microammeter deflections on PCC-2 panel were recorded, which corresponded to the increased optical density of the reaction under study. Since the bursts are uncharacteristic of the instrument properties and admissible estimates under the experimental procedure, it was

suggested that the reason for the observed bursts might be associated with non-trivial fluctuations (pulsed signals) [10, 11]. The general characteristics of pulsed signals is as follows:

— The polarity of pulsed signals corresponded to the decrease of photocurrent magnitude;

— Pulsed signals were observed at any time of the day, including around midnight;

— The duration of bursts was less than one second;

— Bursts were registered under various shielded conditions in the laboratory building coated with duralumin sheets (Antarctica); in ship-board space, multiple-shielded by steel deck grillage (RSV "Akademik Fedorov"); and in a cast-concrete building (AARI, St.Petersburg). The intensity of pulsed signals in Antarctica was considerably higher than in St.Petersburg;

— Pulsed signals have no geographic restrictions, they were recorded both in the Southern and Northern hemispheres, from 70°S (Antarctica, 2000) to 86°N (Arctic, 2000).

Since the receiving unit in the instrument is represented by a photocell, where the photocurrent is recorded with a micro-ammeter, variation recording on the instrument panel was transformed through the recording of photocurrent values in the absence of working substance. Testing of photocurrent measurements was performed in a cast-concrete building (AARI, St. Petersburg) in automatic mode via the COM-port of a PC, using DM3600 digital multimeter. The experiment was supplemented by PCC-2 thermal stabilization 20°C (±1). Uninterrupted power supply to the entire system was ensured by UPS-525 bt. In the course of measurements, abrupt changes of photocurrent in the form of a pulsed signal, in the direction of its decrease, was recorded. Sample registration of pulsed signals in photocurrent is shown in Fig. 2.

The practice of geophysical observation involves methods of testing the effects of artificial electromagnetic interference on the recording systems. Testing may be valid if the experiment is placed away from the metropolis, to minimize the impact of anthropogenic factors. Following these requirements, photocurrent measurements with PCC-2 were conducted at Novolazarevskaya station (Antarctica) in 2004.

3 Checking the integrity of the experiment

At the primary stage of automated measurements, PCC-2 sensitivity to the effects of artificial electromagnetic field (AEMF) was tested. The following instruments were used in the experiment:

1. Coil–to generate an electromagnetic field:
   — Radius of turn: 0.055 m;
   — Coil dimensions: width of turn–0.07 m, diameter–0.11 m;
   — Number of turns: 17;
2. Self-contained DC power source: storage battery 24 V, 75 A×hour;
3. Dropping resistor R_D = 2 Ohms;
4. "Mastech" M-832 digital multimeter;
5. Stationary recording magnetometer at Novolazarevskaya station;
6. PCC-2 photocurrent recording system:
   — PCC-2 microphotocolorimeter;
   — M3850D digital multimeter (with RS-232 cable outlet);
   — Power supply unit for multimeter.

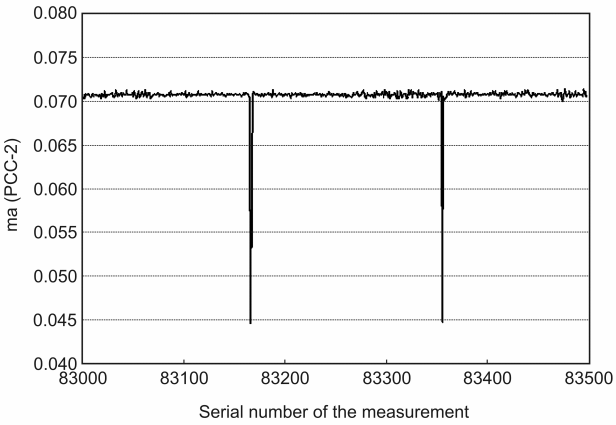

Fig. 2: Sample registration of pulsed signals in photocurrent (PCC-2)

The magnetic field was excited by a pulsed current without a dropping resistor. The protocol of PCC-2 testing effects is provided in Table 1.

3.1 Measurements of the magnetic field generated by the coil

Measurements of the magnetic field induced by coil were conducted with stationary recording magnetometer (SRM) at Novolazarevskaya station. The magnetic field was excited by a pulsed current without a dropping resistor, and by a constant current with a dropping resistor (R = 2 Ohms). Generation of DC magnetic field for longer than 1 second without limiting resistance was not possible, because the strength of current could be as high as 100A against the resistance of coil ~ 0.2 Ohms. For this case, the magnetic field of coil was measured additionally with dropping resistor included in the circuit.

3.2 Measurement results

Since the three-component magnetic variometer can not be considered as a point-source instrument at a distance of 1 me-

Time (GMT)	The distance to the PCC-2	The position of the coil	The mode of influence	Vector magnitude of induction of a magnetic field (nT)
16:16:30	1.5 m	vertical	a single pulse pulse duration ~ 0.2 c	450 – 900
16:17:00	1.5 m	vertical	a single pulse pulse duration ~ 0.2 c	450 – 900
16:17:35	1.5 m	vertical	a single pulse pulse duration ~ 0.2 c	450 – 900
16:18:20	1.5 m	horizontal E-W	a single pulse pulse duration ~ 0.2 c	450 – 900
16:18:30	1.5 m	horizontal E-W	a single pulse pulse duration ~ 0.2 c	450 – 900
16:18:40	1.5 m	horizontal E-W	a single pulse pulse duration ~ 0.2 c	450 – 900
16:19:30	1.5 m	horizontal N-S	a single pulse pulse duration ~ 0.2 c	450 – 900
16:19:45	1.5 m	horizontal N-S	a single pulse pulse duration ~ 0.2 c	450 – 900
16:20:00	1.5 m	horizontal N-S	a single pulse pulse duration ~ 0.2 c	450 – 900
16:20:15	1.5 m	horizontal N-S	a single pulse pulse duration ~ 0.2 c	450 – 900
16:21:30	1.9 m	vertical	a single pulse pulse duration ~ 0.2 c	250 – 500
16:21:40	1.9 m	vertical	a single pulse pulse duration ~ 0.2 c	250 – 500
16:21:45	1.9 m	vertical	a single pulse pulse duration ~ 0.2 c	250 – 500
16:22:00	1.9 m	horizontal E-W	a single pulse pulse duration ~ 0.2 c	250 – 500
16:22:10	1.9 m	horizontal E-W	a single pulse pulse duration ~ 0.2 c	250 – 500
16:22:15	1.9 m	horizontal E-W	a single pulse pulse duration ~ 0.2 c	250 – 500
16:23:00	1.9 m	horizontal N-S	a single pulse pulse duration ~ 0.2 c	250 – 500
16:23:10	1.9 m	horizontal N-S	a single pulse pulse duration ~ 0.2 c	250 – 500
16:23:15	1.9 m	horizontal N-S	a single pulse pulse duration ~ 0.2 c	250 – 500

Table 1: Results of AEMF testing effects on PCC-2.

ter (three sensors located along a straight line, at a distance of 16 cm), the numeric value (module of vector) of magnetic induction in the coil with current could be calculated only approximately, based on the data from three variometers. Variations in 3 components of the magnetic field were registered – D (WE), H (SN), and Z (vertical). Module of the vector T was calculated by the equation:

$$T = \sqrt{\Delta D^2 + \Delta H^2 + \Delta Z^2}. \quad (1)$$

The results of calculations are presented in Table 2.

Given that the value of coil resistance R_{cat} equaled 0.2 Ohms and in dropping resistor 2 Ohms, it could be expected that the magnetic field in the coil carrying current would be ~10 times higher with the switched dropping resistance than without it. However, we did not account for the internal resistance of the battery, which depends both on the current frequency (essential in case of a current pulse) and its strength. Table 2 demonstrates that other conditions being equal, a coil powered by the same battery creates a magnetic field, which is 4–5 times greater under dropping resistor, as compared to the field generated without it. This ratio is probably even more, due to the low frequency of ADC sampling used in measurements of pulsed fields.

As is known, the maximum and minimum values of magnetic induction vector differ two-fold exactly, if measured equidistant from the center of the magnetic dipole. The coil size being ~ 0.1 m, the field of the coil can be regarded with good accuracy as a dipole field, at a distance above 1 m from its center. Thus, when the coil is powered by a battery in a pulsed mode, magnetic induction measured at 1 m off the coil center ranges from 1500 nT (in the plane perpendicular to the coil axis) to 3000 nT (on its axis). The findings of the experiment showed that the photocurrent readings in PCC-2 are not affected by pulsed electromagnetic field with the magnetic component value > 6000 nT, which is greater by 2–3 orders of magnitude than the maximum amplitude of geomagnetic pulsations at 0.1–0.001 Hz frequencies, and several times higher than the intensity of the strongest magnetic storms. The second experiment on the effects on PCC-2 was conducted with high-frequency transmitter (1782 MHz), ACS-1 radiosonde aerological service at Novolazarevskaya station. Characteristics of the transmitter: — Operating frequency 1782±20 MHz — Pulse recurrence frequency 457.5±0.2 Hz — Pulse duration 1 mcs — Transmitter power 2 W / 300 W The distance between the PCC-2 location "geophysicists' premises") and the aerological service was measured with a GPS receiver and made 145±15) m along a straight line. Effects of ACS-1 were estimated through sessions, of 18 minutes total duration, in 2 W and 300 W modes.

Emission series under transmitter power 2W: a) from 22^h 34^m till 22^h at 37 m, at 0° vertical deviation b) from 22^h 37^m till 22^h 39^m, at +1° vertical deviation c) from 22^h 39^m till 22^h 43^m, at +3° vertical deviation, aimed to receive the signal reflected from the adjacent rocks (~ 50 m) and the ice cap (~ 500 m). The signal reflected from the ice cap was fixed on the radar screen. The signal reflected from the rocks was within the measurement error, due to time delay arising from the proximity to the transmitter. Emission series under transmitter power 300W: a) from 22^h 51^m till 22^h 56^m, at 0° vertical deviation b) from 22^h 56^m till 22^h 59^m, at +1° vertical deviation c) from 22^h 59^m till 23^h 02^m, at 1° vertical deviation.

The experimental result proved that photocurrent readings are unaffected by PCC-2 exposure to the high-frequency electromagnetic field.

4 Data analysis and search of driving factor

Data processing and analysis of photocurrent measurements were carried out with "Statistica" software using the following statistical methods:

— Calculation of the parameters of the distribution (standard error, standard deviation, variance);

$$S = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2}. \quad (2)$$

— Spectral (Fourier) analysis (periodograms, estimate of the spectral density), cross-analysis, the value of coherence;

— Identification of the time series model (trend analysis), and analysis of the inadequacy of the model (analysis of residuals), $e_i = (y_i - y_i - hat)$;

— Parabolic polynomial interpolation of the best approximation $y = b_0 + b_1x + b_2x^2 + b_3x^3 + ... + b_nx^n$;

— Selection of the filtering method, use of the moving average (from 3 to 23 points);

— Cross-correlation, correlation factor (r).

During the period of photocurrent measurements with PCC-2 at Novolazarevskaya station, from 24.04.2004 till 01.02.2006, over 20,000 events of pulsed signals were registered. The average daily number of signals comprised ≈300 events, with a minimum of about 80 and a maximum of 580 events. All registered signals are characterized by the polarity in the direction of decreasing photocurrent, 30–50%, on average. The long-period variations of about one year duration may be distinguished in the general distribution pattern of pulsed signals (Fig. 3). The figure also reveals that broad maxima correspond to the end periods of the polar night (July–August). Hence, the number of signals (intensity) does not depend on the influx of solar radiation. In search for the connection of these variations with cosmophysical factors, attention was given to the annual Earth motion along the orbit (the ecliptic). As is known, the equation of time [12] is the sum of two following components. These are the *eccentricity equation* and the *ecliptic inclination equation* (the declination

Time (GMT)	The distance to the SRM	The position of the coil	The mode of influence	Vector magnitude of induction of a magnetic field (nT)
15.07.2006 14:03:10–14:03:16	2 m	horizontal E-W	a single pulse pulse duration ~ 0.2 c	~ 200 nT
15.07.2006 14:06:33–14:06:34	1.5 m	horizontal E-W	a single pulse pulse duration ~ 0.2 c	~ 700 nT
15.07.2006 14:09:42–14:09:51	1 m	vertical	a single pulse pulse duration ~ 0.2 c	~ 2500 nT
20.07.2006 10:43:30–10:44:00	1 m	vertical	constant field	~ 600 nT
20.07.2006 10:51:00–10:51:35	1 m	vertical	constant field	~ 650 nT
20.07.2006 10:53:45–10:54:15	1 m	vertical	constant field	~ 550 nT
20.07.2006 10:56:45–10:57:15	1 m	horizontal E-W	constant field	~ 360 nT

Table 2: Results of AEMF testing effects on SRM.

of the Sun). While the bonding of the connection of the signals and the *ecliptic inclination equation*, we obtained a correlation whose coefficient is close to $r \sim 0.7$. Fig. 3 (a, b) gives the comparision of the numerical values of the daily impulse signals in the photocurrent (KFK-2) and the numerical values of the Sun's declination during 24.04.2004–01.02.2006.

In general distribution of the signals, variations of different duration are traced. Their behavior was identified by comparing the total signal distribution with the indices of solar activity and the total solar radiation (TSI), as well as with fluxes of solar cosmic rays and geomagnetic activity indices. These comparisons revealed that the changes of the daily values of signals best correspond to the SA F10.7 cm index changes and the average daily standard deviation of energy TSI (SD). Standard deviation (SD) shows the variance of the random variable values, with respect to its statistical expectation, i.e., the rate of within-group variability of a given indicator. Comparisons of the series are shown in Figs. 4–5 (a, b). The less pronounced relationship is viewed in case of K-index (Fig. 6) and the SCR fluxes (Fig. 7). Figure 7 demonstrates good matching in the value's trends starting from 425 days (late June 2005).

5 The parallel measurements of the photocurrent (KFK-2) and the fluctuations at the 520 nm wavelength, in the light-emitting diode 565 nm (AvaSpec-2048).

Assuming that the effects in PCC-2 photocurrent were caused by heliophysical impact, similar effects should be expected in the readings of other instruments.

AvaSpec-2048 (www.avantes.com) is a multifunctional fiber optic spectrometer intended for a wide range of studies (Fig. 8). The spectrometer is designed on AvaBench-75 platform with symmetric optical bench (Czerny-Turner). The elemental profile of spectral distribution is read by the operated electronic board and is further transferred from the detector matrix to PC via USB/RS-232.

The task of the second experiment was to conduct "PCC-2 –AvaSPec–2048" parallel measurements referenced to GPS universal time. The measurements were performed from 16.05.05 till 01.11.05, with spatial separation of the instruments up to 5 m distance, in a continuous automatic mode. When processing fluctuations in different LED (565nm) spectral lines, the 520 nm line was selected, where the observed pulsed signals had the same specifics as in PCC-2. Fig. 9 shows an example of the registration. The first estimates of fluctuations comparing the two methods were obtained for the period from 31.05.04 till 08.09.04. Figure 10 shows the temporal comparisons for daily values of bursts in photocurrent FD-7K and energy fluctuations the wavelength 520 nm converted into the average daily standard deviation (SD).

6 Parallel measurements of photocurrent (PCC-2) and fluctuations within the 339.5–346 nm range of the atmosphere zenith (AvaSpec–2048)

Measurements of fluctuations within the 339.5–346 nm range of the atmosphere zenith were conducted with fiber optic spectrometer AvaSpec–2048. The data acquisition chart on spectral zenith observations of solar UV–radiation is presented in Fig. 11.

The measurements were performed during the polar summer, in accordance with the methodology of zenith observations on the ozone content, at the Sun angle $> 5^\circ$. Data were recorded in the files in automatic mode, with a sampling interval of 2–3 seconds. Observations were accompanied by time corrections from GPS. The initial phase of observations aimed at the search of non-instrumental fluctuations at the full range of frequencies, within 297–780 nm range. At 0.3 nm resolution of diffraction grating, more than 1,300 spectrum lines were analyzed. To conduct parallel measurements with

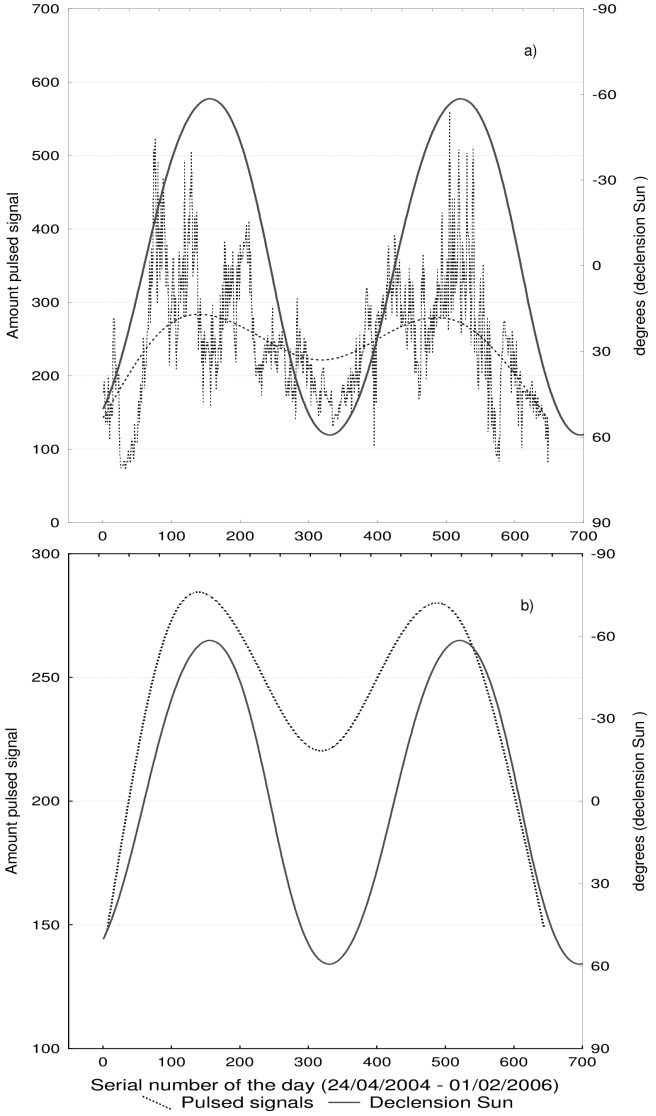

Fig. 3: the comparison of the distributions of the daily numerical values of the impulse signals of KFK-2 and the general number polynomial b) and the ecliptic inclination (the declination of the Sun) during 24.04.2004–01.02.2006 (Novolazarevskaya station).

the spectrometer and PCC-2, four ranges of the atmosphere zenith spectrum were selected (303–305 nm, 331–332.5 nm, 329.5–334 nm, 339.5–346 nm), for which the standard deviation of energy (SD_E) exceeded the instrumental fluctuations by an order, or above [13, 14]. Figure 12 shows a sample recording of fluctuations in the range of 339.5–346 nm. The profile demonstrates bipolar fluctuations reaching 339.5 nm and 346 nm levels, measured in the center of 342.5-nm frequency range.

The energy estimates (eV/photon) of pulsed signals were defined. For example, according to the formula:

$$\text{photon energy} \quad E(\lambda) = \frac{h\,c}{\lambda_e}\,, \tag{3}$$

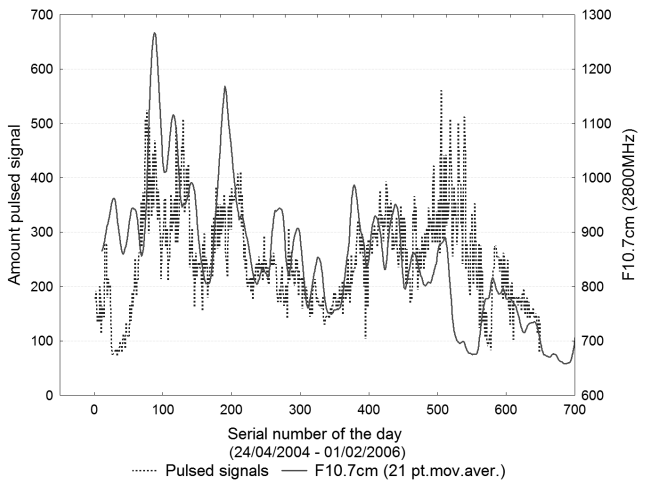

Fig. 4: Comparison of temporal changes in solar activity index F10.7 cm (2800 MHz) with distribution of the daily values of pulsed signals in PCC-2, for the period 24.04.2004–01.02.2006 (Novolazarevskaya station).

where, h = Planck's constant 6.62606876×10^{-34}; c = velocity of light, 2.998×10^{8} m/s; λ = wavelength in meters.

The average estimates of pulsed signals of energy within the 339.5–346 nm range were as follows:

$$E_{min}\,(346\text{nm}) = 3.583\,(\text{eV/photons}),$$
$$E_{mean}\,(342.5\text{nm}) = 3.619\,(\text{eV/photons}),$$
$$E_{max}\,(339.5\text{nm}) = 3.652\,(\text{eV/photons}).$$

Comparison of fluctuations within the tested ranges with pulsed signals in PCC-2 showed an ambiguous correlation. The most consistent changes in PCC-2 pulsed signals were observed within a 339.5–346 nm range. Comparison of the series for the period from 25.09.07 till 17.12.07 illustrates this example in Fig. 13.

7 Prognostic functionalities of the observed effects

In addition to the obtained results, the general distribution pattern was detected in the daily values of pulsed signals measured with PCC-2, which corresponded to ≈ 300-day cycle. This period was revealed through the comparison of the annual intervals of the general series, with a difference of about two months. The second interval was compared to the first against the difference minus ≈ 60 days from the start of the first interval. For example, the top graph in Figure 14 (a, b) shows the intervals comparison for 24.04.2004–23.04.2005 and 21.02.2005–01.02.2006. On the bottom figure, the cross-correlation function of two series by logs, with corresponding correlation factors, is shown. It can be seen that the maximum correlation value reaches $r \sim 0.7$. Assuming all the above relationships between pulsed signals and variations of the SA index and TSI, a 300-day cycle was also detected in the F10.7 cm index and TSI. A comparison of F10.7 cm and TSI distribution patterns within the annual intervals is provided in

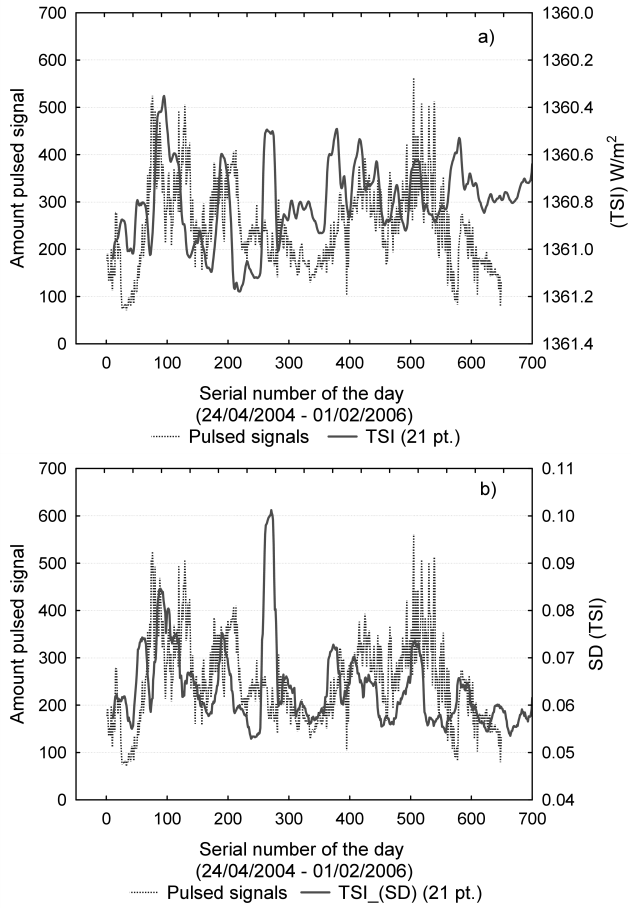

Fig. 5: Comparison of temporal changes in energy TSI, and b) — Comparison of the daily average standard deviation TSI (SD) with distribution of the daily values of pulsed signals in PCC-2, for the period from 24.04.2004 till 01.02.2006 (Novolazarevskaya station).

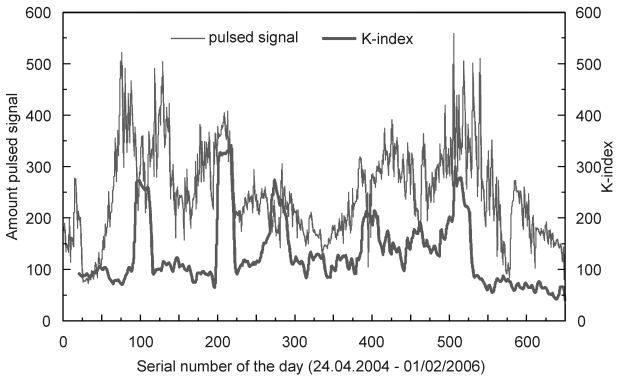

Fig. 6: Comparison of temporal changes in K-index variations with distribution of the daily values of pulsed signals in PCC-2, for the period from 24.04.2004 till 01.02.2006 (Novolazarevskaya station).

Fig. 15 (a, b): 24.04.2004–24.04.2005 and 24.02.2005–01.02.2006. Figure 15 (a) indicates the matching of F10.7 cm index variation in phase opposition of variations.

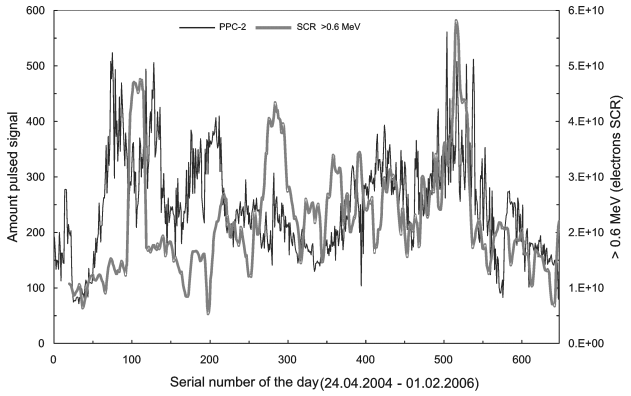

Fig. 7: Comparison of the temporal changes in SCR electron variations (> 0.6 MeV) with distribution of the daily values of pulsed signals in PCC-2, for the period from 24.04.2004 till 01.02.2006 (Novolazarevskaya station).

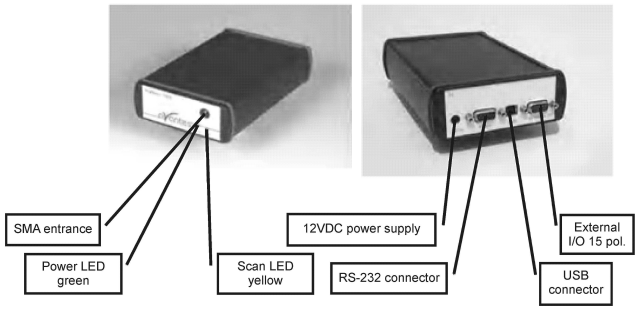

Fig. 8: Spectrometer AvaSpec-2048 (www.avantes.com).

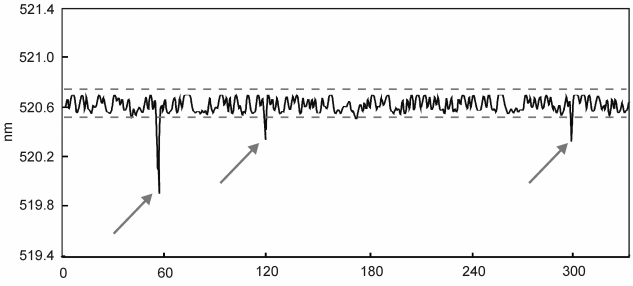

Fig. 9: Sample recording of pulsed signals at 520-nm frequency in 565-nm LED spectrum (AvaSpec–2048) (Novolazarevskaya station).

Comparison is presented upon the unfiltered values (without leveling). Unlike F10.7 cm, the daily average standard deviation of the energy TSI reveal matching of variations in the same phase character but if applying the moving average filter to 21 pts. The maximum amplitude in SD (TSI) is associated with a solar flare.

As regards the contribution of geomagnetic factors to the 300-day period, its manifestation was traced in the interplanetary magnetic field component (By). Figure 16 shows the comparison of By values for the intervals 01.01.2003–

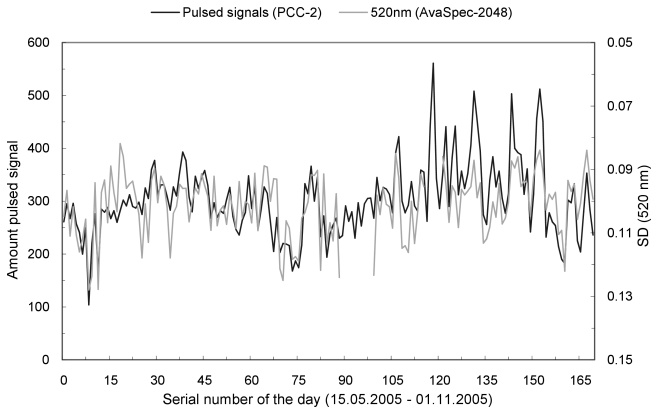

Fig. 10: Comparison of daily values of bursts in photocurrent with the average daily standard deviation (SD) of fluctuations at the 520 nm wavelength (AvaSpec–2048), for the period from 15.05.05 to 01.11.05 (Novolazarevskaya station).

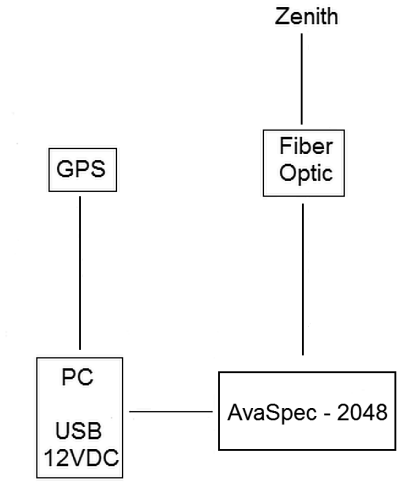

Fig. 11: Data acquisition chart on spectral zenith observations of solar UV–radiation in the atmosphere zenith (Novolazarevskaya station).

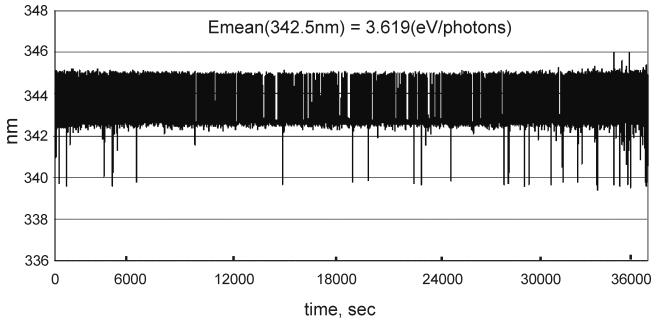

Fig. 12: Sample registration of fluctuations in 339.5–346 nm range (AvaSpec–2048), at clear atmosphere zenith, from $07^h\ 00^m$ till $17^h\ 00^m$ (09.03.2005, Novolazarevskaya station).

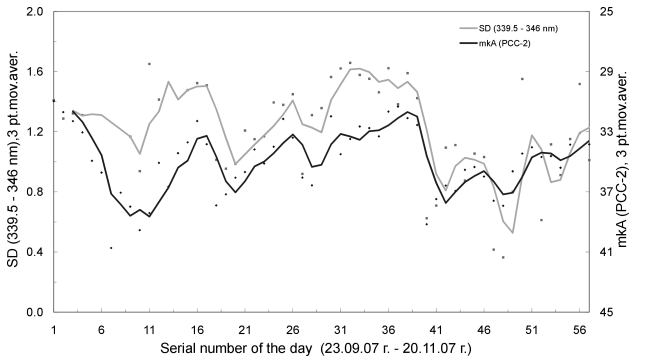

Fig. 13: Comparison of the daily average standard deviation of fluctuations in energy (SD_E) within the 339.5-346 nm range (AvaSpec-2048) in the atmosphere zenith and PCC-2 pulsed signals, for the period from 23.09.07 to 20.11.07 (Novolazarevskaya station).

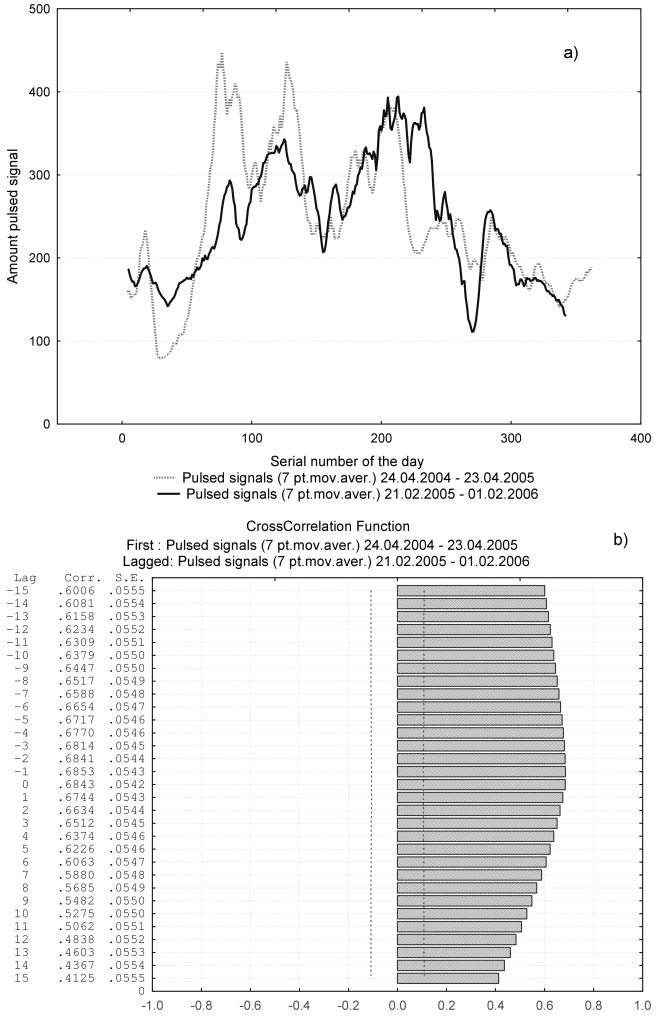

Fig. 14: Comparison of PCC-2 daily values of pulsed signals within the intervals 24.04.2004–24.04.2005 (a) and 21.02.2005–01.02.2006 (b) (Novolazarevskaya station).

01.01.2004 and 27.10.2003–27.10.2004. The identity in variations and the phase convergence of the series, as demonstrated in Figure 16, is the most indicative of the existence of the 300-day cycle. As in the case with F10.7 cm, unfil-

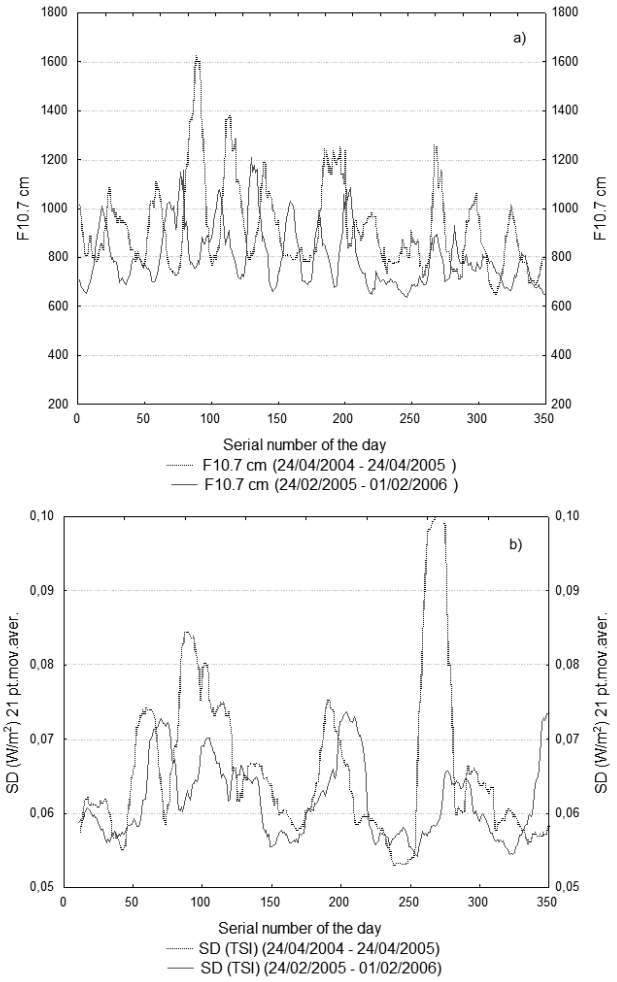

Fig. 15: Comparison of F10.7 cm and TSI distribution patterns within the intervals 24.04.2004–24.04.2005 (a) and 24.02.2005–01.02.2006 (b).

tered series were compared. This convincing fact is not yet explained through the known mechanisms of solar-terrestrial relationships.

Thus, pulsed signals bear prognostic characteristics associated, in our opinion, with the unknown heliophysical factor.

Similar results were obtained under extensive laboratory experiments conducted by Sergey Korotaev at Geoelectromagnetic Research Institute RAS (Troitsk, Moscow Region, Russia) [6].

His publication considers the phenomenon of non-locality and correlation of isolated dissipative processes, as well as description of the experiment and the results of investigations. Using two types of detectors based on the link between the entropy and the potential barrier height U, reliable correlations were established between: lead variations of dark current and temporal events in the atmospheric pressure (69, 73 days), variations of geomagnetic activity (Dst-index) (33 days) and variations of the F10.7 cm index (42 days). Based

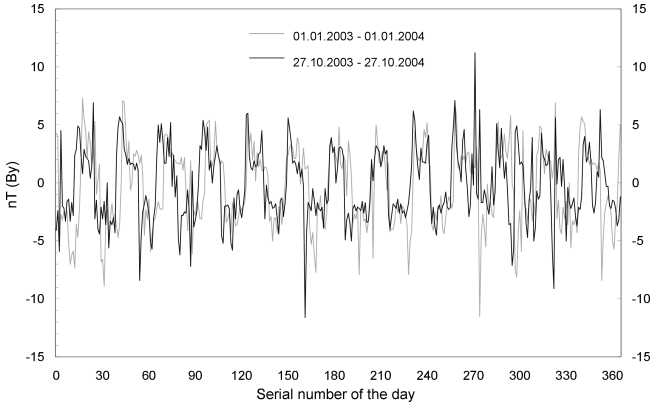

Fig. 16: Comparison of variations of the interplanetary magnetic field component By, by intervals: 01.01.2003–01.01.2004 and 27.10.2003–27.10.2004.

on the established correlations between the detectors and SA, the author established a temporal range of the lead coupling, from 42 to 280 days.

8 Conclusion

In our assumptions on the source of pulsed signals, the impact of cosmic particles fluxes and their secondary radiation (protons and products of their interaction with atmospheric nuclei — *positrons, muons, pi-mesons, K-mesons, electron pairs, gamma ray quanta, atmospheric neutrinos,* etc.) should be first considered. In determining the relation of the observed pulsed effects to the secondary cosmic rays, a comparison of distribution of pulse signals with variations of atmospheric cosmic rays of geomagnetic, solar and galactic origin would be quite sufficient. Assuming that similar pattern in temporal variations of the compared series would persist for a long time interval (within months), the penetrating component should be identified, since pulses are observed in shielded conditions. The penetrating component of cosmic rays can be defined by the type of interaction of cosmic rays with the substance.

For instance, among those well-known are: the nuclear-active component, soft component of secondary cosmic rays, electron-photon showers and the penetrating component of the secondary radiation — muons and neutrinos. Of the above mentioned, muons are the most likely source, as the muon flux represents a penetrating component and, against a relatively moderate energy power (~10 GeV), can easily enter the atmosphere and penetrate the shielding conditions of PCC-2. However, this version contradicts with the significant differences between the diurnal statistics of muons and neutrinos and pulsed signals. E.g., according to the observations in NT-200 neutrino telescope located at a depth of 1100–1200 m in the Baikal Lake, the daily number of atmospheric muons reaches ≈1,000,000 and for atmospheric neutrinos – one oc-

currence in two days, on average.*

The hypothetical factor of heliophysical origin that causes simultaneous effects in photocurrent, the LED spectrum, and the zenith spectrum of the atmosphere, remains unknown. On the one hand, a good coincidence between the number of signals and variations in F10.7cm and TSI is observed, which may be regarded as conclusive indication of their solar origin. On the other hand, signals are registered, regardless of the shielding conditions, which is indicative of a high penetration capacity of the heliophysical factor. Further, no reliable evidence of a link between pulsed signals, cosmic ray fluxes and geomagnetic activity is revealed. It is as well obvious that the known mechanisms of the solar-terrestrial relationships do not represent direct implications of the effects observed in the experiment. An especially characteristic indicator of the signal intensity is the statistical correlation with the declination of the Sun, i.e., the position of the Earth on the orbit (the ecliptic). We do not exclude the assumption about the directed impact of this "hypothetical" factor in the ecliptic plane.

According to the results demonstrated in figures 14–16, it is evident that the studied pulsed signals in photocurrent have 60–65 days lead time, on average, compared to the solar events. If such pattern is scaled against the 11-year cycle of the solar activity (SA), we should expect that solar variations, recurrent in 300 days, would return to the starting point in 5.5 years (in our case, the measurement starting point was 24.04.2004), which makes approximately a half of the 11-year SA cycle. This presents a point of interest. The solar cycles are known to have a progression: 11-year, 22-year, 44-year, 88-year, etc. According to our results, we can not exclude the possibility of declining values of the SA cycles, down to 5.5 years, or less. Possibly, the 300-day cycle refers to the initial cycles in this progression. Its physical component may be determined by the processes occurring in the central zone of the Sun.

Submitted on: April 26, 2013 / Accepted on: May 06, 2013

References

1. Chizhevsky A.L. Cosmic pulse of life. Mysl', Moscow, 1995.
2. Kozyrev N.A. Selected Papers. Leningrad Univ. Publ., Leningrad, 1991.
3. Lavrent'ev M.M., and others. On the registration of the true position of the Sun. *Doklady AN USSR*, 1990, v.315, no.2, 368–370.
4. Shnol' S.E., and others. Regular variation of the fine structure of statistical distributions as a consequence of cosmophysical agents. *Physics-Uspekhi*, 2000, v.170, no.2, 214–218.
5. Parkhomov A.G., Maklyaev E.F. Research on rhythms and fluctuations through long-term measurement of radioactivity, crystal frequency, semiconductor noise, temperature and atmosphere pressure. *Fizicheskaya Mysl' Rossii* (*Russian Physics Reports*), 2005, no.1, 1–12.
6. Korotayev S.M. Heliogeophysical nonlocality effects — shadows of future in the present. *Quantum Magic*, 2004, v.1, issue 2, 2219–2240.
7. Kondratyev K.Y., Nikol'skiy G.A. Impact of solar activity on structure components of the Earth. *Issledovanie Zemli iz Kosmosa*, 2005, no.3, 1–10.
8. Sizov A.D. Current fluctuations in the Wheatstone bridge. Possible cosmophysical correlations. *Biofizika* (*Biophysics*), 1998, v.43, issue 4, 726–729.
9. Gorshkov E.S., Shapovalov S.N., Sokolovskiy V.V., Troshichev O.A. On the reaction rate of the unitiol oxidation by nitrite ion. *Biofizika* (*Biophysics*), 2000, v.45, issue 4, 631–635.
10. Shapovalov S.N., Gorshkov E.S., Troshichev O.A. Cosmophysical effects observed in impulses of the microphotocolorimeter current. *Biofizika* (*Biophysics*), 2004, v.49, Suppl. 1, 119–121.
11. Shapovalov S.N., Troshichev O.A., Povazhny V.I. Study of short time pulses in the photoelectric effect under conditions of Antarctica (station Novolazarevskaya). *T2-8 Heliosphere Impact on Geospace IPY Oslo Science Conference*, 08–12 June 2010.
12. The Astronomical Calendar (permanent part). Nauka, Moscow, 1981.
13. Troshichev O.A., Shapovalov S.N., Lozovsky V.T. Study of pulsed signals in UV spectra lines of free atmosphere above Novolazarevskaya station (Antarctica): effect of the solar irradiance? *37th COSPAR Scientific Assembly*, 13–20 July 2008, Montreal, Canada, p.322.
14. Shapovalov S.N., Troshichev O.A. Study of pulsed energy fluctuations and solar UV variations by data of spectral measurements in zenith of free atmosphere at Novolazarevskaya station (Antarctica). *39th COSPAR Scientic Assembly*, 14–22 July 2012, Mysore, India (C2.3-0009-12).

*http://nuclphys.sinp.msu.ru/neutrino/newtrino_s/baik.htm

Is Space-Time Curved?

Benjamin Prather
E-mail: benjamin.prather@gmail.com

This paper considers the possibility of a teleparallel approximation of general relativity where the underlying space-time of a compact massive source is related to the isotropic coordinate chart rather than the geometric chart. This results in a 20 percent reduction of the expected shadow radius of compact objects. The observation of the shadow radius of Sagittarius A* should be possible in the near future using VLBI. The theoretical reduction is within the uncertainty of the expected shadow radius, however any observation less than a critical radius would indicate that gravity is not the result of space-time curvature alone. If space-time curvature does not act alone it is simpler to adopt the teleparallel view, with the tetrad field representing the index of refraction of the required material field in a flat space-time.

Introduction

General relativity is highly successful in explaining the first order corrective terms to Newtonian gravity observed in the classical solar system test known at the time of its proposal. Further, it has predicted higher order effects not originally anticipated such as the orbital decay of binary pulsars. Any competing theory of gravity must agree with general relativity in these predictions. The bounds on these measurements have significantly improved since the introduction of general relativity [1].

The central tenant of general relativity is that gravity is a pseudo-force due to the curvature of space-time. This produces a theory lacking an absolute sense of parallelism. General relativity has been expressed as a teleparallel theory, thus restoring absolute parallelism [2].

The teleparallel equivalent of general relativity allows the curvature of a metric to be rephrased as contorsion in a flat space-time due to a tetrad field [2]. The geodesic equation becomes non-inertial forces as a result of the variation in the local index of refraction and motion the tetrad field represents.

This paper considers the implications of a teleparallel theory of gravity where the underlying space-time corresponds to a flattened version of the isotropic solutions rather than the usual geometric coordinates. This non-inertial flattening process produces pseudo-forces, which are taken to be actual forces due to the presence of a material field.

Globally, space-time is likely to be closer to a DeSitter space-time than the Minkowski space-time used in the limiting behaviour here. In this sense, space-time is demonstrably curved. The issue here is the local nature of space-time in the presence of strong gravitational fields.

1 Is space-time curved?

Despite its broad empirical success, and lack of any viable alternatives, general relativity continues to generate detractors who raise philosophical objections to its core propositions. These detractors, near or beyond the fringe of science, often lack the mathematical knowledge needed to properly discuss general relativity in a rigorous setting. Indeed, many of these objections stem from a rejection of the abstract mathematics required for general relativity or perceived errors in general relativity arising from subtle misunderstandings of these advanced notions.

The descendants of neo-Kantianism assert that space-time curvature caused by matter and energy is impossible, since matter and energy already require the concepts of space and time. A Galilean space-time is also claimed by these critics to be necessary to form an understanding of the world [3].

As Lie groups, however, the Poincaré group is equally descriptive as the Galilean group. These differences in symmetries can be empirically measured, strongly favouring a Minkowskian space-time over a Galilean space-time. In both geometries it is almost always helpful to select a convenient fixed frame to work within. General relativity, in its usual presentation, breaks this Lie symmetry globally.

General relativity can also be expressed as a teleparallel theory, restoring absolute parallelism by replacing the curvature of space with an embedded tetrad field. Tetrad fields can be viewed as representing the flow and refractive properties of a Lorentzian aether.

In classical fluid mechanical one can use a Lagrangian reference frame co-moving with a fluid or an inertial Eulerian reference frame. In a relativistic aether, using the Levi-Civita connection produces the Lagrangian description while using the Weitzenböck connection produces the Eulerian description.

The geodesic equation then becomes changing speed due to an index of refraction, bending due to Huygens' principle and frame dragging due to advection. In the teleparallel equivalent of general relativity this tetrad field exists as an independent structure. This can be viewed as a flowing index of refraction emerging in the absence of a refractive medium.

This theory can be bashed into a flat model using a non-inertial transformation. The use of a non-inertial reference

frame introduces pseudo-forces to the equations of motion. Interpreting these forces as originating from a material field creates a flat theory of gravity while simultaneously providing a material medium responsible for the tetrad field.

It is in this sense that the question is raised, is space-time curved?

2 Flat teleparallel approximation

In general relativity, the gravity of a compact, spherically symmetric, uncharged, acceleration-free and isolated mass generates can be described by the well known Schwarzschild metric in spherical coordinates.

$$r_s = \frac{2GM}{c^2}\,, \tag{1}$$

$$\mathbf{g}_{ij} = \mathrm{diag}\left(\left(1-\frac{r_s}{r}\right)c^2,\ \left(1-\frac{r_s}{r}\right)^{-1},\ r^2,\ r^2\sin^2\theta\right). \tag{2}$$

This solution implies that the speed of light depends on the angle of inclination of the trajectory relative to the coordinate chart. It is possible to transform the radial component to a new chart where the speed of light is isotropic [4].

$$r = r'\left(1+\frac{r_s}{4r'}\right), \qquad r' = r\left(\frac{1}{2}-\frac{r_s}{4r}+\sqrt{\frac{1}{4}\left(1-\frac{r_s}{r}\right)}\right), \tag{3}$$

$$\mathbf{g}_{ij} = \left(1+\frac{r_s}{4r'}\right)^4 \mathrm{diag}\left(\frac{(4r'-r_s)^2}{(4r'+r_s)^6}\,c^2,\ 1,\ r^2,\ r^2\sin^2\theta\right). \tag{4}$$

A flat teleparallel approximation of general relativity can be made by eliminating the $\left(1+\frac{r_s}{4r'}\right)^4$ coefficient. This results in a flat space-time with an index of refraction.

$$\mathbf{g}_{ij} = \mathrm{diag}\left(\frac{(4r'-r_s)^2}{(4r'+r_s)^6}\,c^2,\ 1,\ r^2,\ r^2\sin^2\theta\right). \tag{5}$$

Determining the pseudo-forces caused by this non-inertial transformation, much less the fields needed to generate them, is beyond the scope of this paper. All that is of interest here is that a model should exist with this geometry in the limit of the Schwarzschild metric, and has an event horizon a quarter the size of general relativity.

Bashing the Schwarzschild metric into a flat teleparallel theory may be a convenient way to get a model that agrees with observation, but is a very ad hoc way to approach the problem. A far better approach would be to build a teleparallel theory from the ground up based on first principles. Once the numerous obstacles are overcome, any resulting theory will agree with general relativity in the weak field limit.

This would require differences in the strong field limit to distinguish between theories. Given the significant change in event horizon radius, the optical shadow radius of a compact object should provide a useful parameter to compare potential theories in the strong field limit.

3 Optical shadow, General Relativity

An image showing the neighbourhood of the singularity, including the event horizon, photon sphere and optical shadow is given in Figure 1.

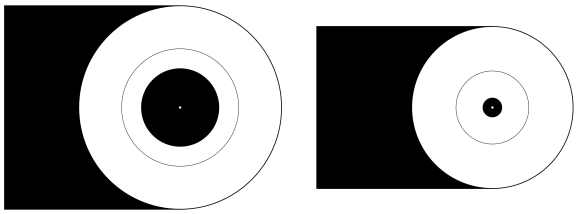

Fig. 1: Singularity Neighbourhood. The neighbourhood of a compact gravitational source is shown, with the Schwarzschild solution to the left and the flattened version to the right. The central black circle represents the event horizon, with a white circle showing the location of the singularity. This is surrounded by a thin circle representing the photon sphere. The outermost circle represents the optical shadow of the black hole, which is shown extending to the left using a tangent line approximation.

The depicted shadow region is a right circular cylinder with the singularity on its axis. The optical radius can be defined as the largest radius such that no unbound trajectory can have both an infinite length within the depicted shadow region and avoid the event horizon. In general relativity, the optical radius is $r_{crit} = \frac{3}{2}\sqrt{3}\,r_s$ [5].

4 Optical shadow, flat teleparallel approximation

In the flat teleparallel approximation, the dynamics are expected to be identical except for a rescaling of the radius near the singularity.

Almost all of the shadowing effect occurs near the singularity. An estimate of the asymptotic trajectory can be made using a tangent line to a circle about the singularity with a radius of r_{crit}. The radius of this circle is transformed by (3).

$$r'_{crit} = r_{crit}\left(\frac{1}{2}-\frac{1}{6\sqrt{3}}+\sqrt{\frac{1}{4}\left(1-\frac{2}{3\sqrt{3}}\right)}\right) \approx 0.8\,r_{crit}\,. \tag{6}$$

This reduction in radius is accomplished by the forces generated by a material component of the gravitational field.

5 VLBI measurements

Expected advances in submillimetre very long baseline interferometry are expected to be able to soon resolve the optical shadow of the compact radio source Sagittarius A*, based on the size expected by general relativity [5]. The factor of 0.8 is close enough to unity that the flat teleparallel approximation should produce a visible shadow under similar assumptions.

Observing the optical shadow is confounded by several known issues, much less measuring the radius. The optical

properties of the medium surrounding Sagittarius A*, the 20 percent uncertainty in the mass of and distance to Sagittarius A* and the 10 percent uncertainty introduced by the dependency on the optical shadow on the rotation of Sagittarius A* are three significant issues [5].

A successful imaging of the shadow could help to determine some of these uncertainties, allowing a determination to be made between general relativity and any potential teleparallel theory including a material gravitational field.

While other effects can account for an optical shadow larger than r_{crit} within general relativity, an optical shadow less than $r_{crit} = 30 \pm 7\mu_{as}$ cannot be reconciled with general relativity [5].

6 Conclusion

Given that the event horizon in the flat teleparallel approximation is a quarter of that predicted by general relativity, the reduction in optical shadow of 20 percent is a disappointingly small change. This is less than the expected uncertainty in the optical shadow of Sagittarius A* due to its uncertain mass and possible rotation.

This also means that the same assumptions for observing the shadow expected for general relativity using VLBI can be applied to the flat teleparallel approximation. Such measurements can be expected on the order of years, not centuries.

This reduction would vary for different models of the material gravitational field, possibly resulting in a smaller optical radius. This would confound the ability to observe the optical shadow but simplify the ability to distinguish the predictions of general relativity and the model in question.

While other effects can account for an optical shadow larger than r_{crit} within general relativity, an optical shadow less than r_{crit} would indicate that gravity is not determined by space-time curvature alone.

The teleparallel equivalent of general relativity phrases the effects of gravity as due to an index of refraction in a flat space-time. If this is not acting alone, it is simpler to view this index of refraction as a property of the material field required to explain the super compact optical shadow.

If a super compact optical shadow is demonstrated, space-time curvature should then be abandoned in favour of a material, refractive gravitational field in a flat or DeSitter space-time.

Submitted on June 19, 2013 / Accepted on June 20, 2013

References

1. Will C. M. 2006 Confrontation between GR and experiment. *Living Rev. Relativity*, 2006, v.9(3), (cited on 6/5/2013: http://relativity.livingreviews.org/Articles/lrr-2006-3/)
2. Aldrovandi R., Pereira J.G. An Introduction to Teleparallel Gravity. Springer, New York, 2013
3. D.H. et al. 2013 Criticism of the theory of relativity. *Wikipedia* (cited on 6/5/2013, oldid=551162068)
4. Buchdahl H. A. Isotropic coordinates and Schwarzschild metric. *Int. J. Theor. Phys.*, 1985, v.24, 731–739.
5. Falcke H., Melia F., Agol E., Viewing the Shadow of the Black Hole at the Galactic Center. *Astrophys. J.*, 2000, v.528, L13–L16.

LETTERS TO PROGRESS IN PHYSICS

LETTERS TO PROGRESS IN PHYSICS

Comment on N. A. Kozyrev's "Possibility of Experimental Study of the Properties of Time"

Joseph C. Hafele

Retired Physicist; Home Office: 618 S. 24th St., Laramie, WY, USA
E-mail: cahafele@bresnan.net

More than 60 years ago, N. A. Kozyrev predicted the need for a second universal velocity, one that is associated with rotational motion, in addition to the well-known first universal velocity, the velocity of light, which is associated with linear motion. Kozyrev predicted that there should be additional forces which act along the axis of rotation and are on the order of 10^{-4} or 10^{-5} of the applied forces. For the neoclassical causal theory (Hafele J. C. *Zelm. Journ.*, 2012, v. 5, 134), the values for the ratios for the Moon are in order of magnitude consistent with Kozyrev's predicted ratios.

The neoclassical causal version for Newtonian gravitational theory requires a hypothetical induction field F_λ and a corresponding induction speed v_k [1]. The purpose of this letter is to indicate that, more than 60 years ago, N. A. Kozyrev developed a similar concept for rotational motion in classical Newtonian theory [2].

The experimental verification of Kozyrev's theoretical concepts started in the winter of 1951–1952. Kozyrev stipulates: 1) the velocity of light is a universal velocity that is to be associated with linear motion, and 2) there should be a second universal velocity that is associated with rotational motion. In his notation, c_1 is the known speed of light, and c_2 is an unknown rotational universal speed.

The following is a direct quote from page 199 of Kozyrev's article [2]:

> "Now, utilizing the Plank constant in any scalar universal constant, it is necessary to obtain a value having the dimensionality of velocity. It is easy to establish that the expression
>
> $$c_2 = \frac{\alpha e^2}{h} = \alpha \times 350 \text{ km/sec} \qquad (7)$$
>
> comprises a unique combination of this type. Here e equals the charge of an elementary particle and α equals a certain dimensionless factor. Then, based on (6), at $u = 100$ m/s, the additional forces will be of the order of 10^{-4} or 10^{-5} (at a considerable α-value) from the applied forces."

Kozyrev defines u to be the linear velocity of the rotating object. He finds that the value for $\alpha \cong 2$, and the value for $c_2 \cong 7{\times}10^5$ m/s $\cong 2.3{\times}10^{-3} c_1$ [2, p. 203]. He predicts that a small additional force is proportional to $u/c_2 \cong 100/7{\times}10^5 \cong$ $\cong 1.4{\times}10^{-4}$ [2, p.198], which is the basis for his "of the order of 10^{-4} or 10^{-5}". (The numerical value for e^2/h, 350 km/s, is calculated according to CGS system of units.)

For the neoclassical causal theory, the ratio of the transverse to the radial field for the NEAR flyby at perigee, $g_\text{trt}/g_\text{r} \cong 4{\times}10^{-6}$ [1, p.169]. A better comparison with Kozyrev's theory is obtained by using the case for the Moon, where the orbital motion is nearly circular [1, p.172]. Let v_co be the orbital speed for an equivalent circular orbit. Then

$$v_\text{co} \cong 1{\times}10^3 \text{ m/s}, \qquad \frac{v_\text{co}}{c_2} \cong 1.4{\times}10^{-3},$$

which is within one order of magnitude of Kozyrev's u/c_2. Let $\langle g_\text{trt} \rangle$ be the RMS average value for the time-retarded transverse field for the Moon, and let $\langle F_\lambda \rangle$ be the RMS average value for the induction field. Then

$$\langle g_\text{trt} \rangle \cong 1{\times}10^{-11} \text{ m/s}, \quad \langle F_\lambda \rangle \cong 1.4{\times}10^{-14} \text{ m/s},$$

$$\frac{\langle F_\lambda \rangle}{\langle g_\text{trt} \rangle} \cong 1.4{\times}10^{-3}.$$

These results show that the relative ratios for the secondary fields are close to the same order of magnitude as they are for Kozyrev's theory.

More than 60 years ago, N. A. Kozyrev could not have known about recently discovered flyby anomalies and a lunar orbit anomaly, but he did have an uncanny insight that has now been brought to fruition. If more attention had been paid to Kozyrev's theory, it may have preempted the neoclassical causal theory. It may also be helpful in designing a ground-based instrument for detecting the Earth's time-retarded transverse gravitational field.

Submitted on: April 10, 2013 / Accepted on April 24, 2013

References

1. Hafele J. C. Earth flyby anomalies explained by a time-retarded causal version of Newtonian gravitational theory. *The Abraham Zelmanov Journal*, 2012, v. 5, 134–187.
2. Kozyrev N. A. Possibility of experimental study of the properties of time. *The Abraham Zelmanov Journal*, 2012, v. 5, 188–220.

LETTERS TO PROGRESS IN PHYSICS

Commentary Relative to the Emission Spectrum of the Solar Atmosphere: Further Evidence for a Distinct Solar Surface

Pierre-Marie Robitaille

Department of Radiology, The Ohio State University, 395 W. 12th Ave, Columbus, Ohio 43210, USA.
robitaille.1@osu.edu

The chromosphere and corona of the Sun represent tenuous regions which are characterized by numerous optically thin emission lines in the ultraviolet and X-ray bands. When observed from the center of the solar disk outward, these emission lines experience modest brightening as the limb is approached. The intensity of many ultraviolet and X-ray emission lines nearly doubles when observation is extended just beyond the edge of the disk. These findings indicate that the solar body is opaque in this frequency range and that an approximately two fold greater region of the solar atmosphere is being sampled outside the limb. These observations provide strong support for the presence of a distinct solar surface. Therefore, the behavior of the emission lines in this frequency range constitutes the twenty fifth line of evidence that the Sun is comprised of condensed matter.

Every body has a surface.
St. Thomas Aquinas [1]

Observationally, the chromosphere of the Sun represents a rarefied region located immediately above the solar surface [2–5]. In 1877, Father Angelo Secchi described the chromosphere in detail including, most notably, a description of its spicules [6, p. 31-36]. For just a few seconds prior to and following the onset of totality during solar eclipses, the "flash" emission spectrum of the chromosphere can be detected. Typically, such studies focus on the visible and ultraviolet regions of the electromagnetic spectrum.

The existence of the visible "flash" spectrum has been known since the early days of spectral analysis. In fact, the famous D3 line, first observed in a prominence during an eclipse, would lead to the discovery of helium on the Sun by Pierre Jules César Janssen and Joseph Norman Lockyer [7, 8]. Since then, great attention has been given to identifying the lines which are contained within the flash spectrum of the chromosphere, particularly through the efforts of astronomers like John Evershed [9, 10] and Donald Menzel [11, 12]. In 1909, George Ellery Hale and Walter Adams photographed the flash spectrum outside of eclipse conditions, opening up new avenues for the study of the chromosphere [13, 14]. Today, spectroscopic emission lines in the visible spectrum of the chromosphere and corona continue to be relevant and spectacular images of the solar atmosphere have now been obtained using spectroscopic lines from highly ionized iron (e.g. FeX–FeXIV) [15–18].

Photographing the chromosphere is slightly more complex in the ultraviolet range, since UV light is absorbed by the Earth's atmosphere. As a result, that spectral region of the flash spectrum could not be sampled until the launch of scientific rockets after World War II [3, p. 180]. In 1946, while at the U.S. Naval Research Laboratory, Baum, Johnson, Oberly, Rockwood, Strain and Tousey [19] obtained the first measurements of the Sun's ultraviolet spectrum using a V2 rocket. A flurry of activity in this area soon followed [20–25] and the ultraviolet spectrum of the Sun has now become a field of great scientific interest [26–28].

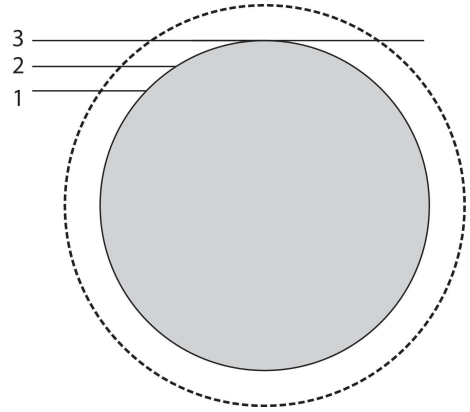

Fig. 1: Schematic representation of path lengths present when the outer atmosphere (area outlined by dashes) of the Sun (body in gray) is viewed from the Earth. Paths 1 and 2 terminate on the solar surface. Just beyond the limb, path 3 samples the front and back side of the solar atmosphere, resulting in a two fold increase in line intensity. This figure is an adaptation based on Fig. 2.4 in [28].

An elementary observation constitutes the focus of this work: the intensity of ultraviolet and X-ray emission lines increases dramatically, as observations are moved from the center of the solar disk to the limb of the Sun. The problem is illustrated in Figure 1. Harold Zirin describes the associated findings as follows: *"The case in the UV is different, because the spectrum lines are optically thin. Therefore one*

would expect limb brightening even in the absence of temperature increase, simply due to the secant increase of path length. Although the intensity doubles at the limb, where we see the back side, the limb brightening inside the limb is minimal . . . Similarly, X-ray images show limb brightening simply due to increased path length." [29]. This situation is observed both in the ultraviolet and in the X-ray spectrum of the Sun which sample processes in the chromosphere and the corona [28, p. 38-39]. An exquisite image of this effect has been published [28, p. 38].

Though this simple observation appears almost trivial as a source of scientific comment, it nonetheless demands attention; for it provides strong evidence that the body of the Sun is not gaseous in nature. If the Sun is gaseous, then these effects should not be visible as sampling extends beyond the solar limb. As such, this observation constitutes the twenty fifth line of evidence that the Sun is comprised of condensed matter (see [30–32] and references therein for the others).

Dedication

This work is dedicated to Amir Abduljalil in recognition of his many years of faithful scientific collaboration throughout my career in magnetic resonance imaging, and for his undying service to The Ohio State University relative to the design, assembly, and operation of the world's first ultra high field magnetic resonance imaging system [33].

Submitted on: February 22, 2013 / Accepted on: February 24, 2013
First published online on: May 13, 2013

References

1. Aquinas T. Summa Theologica — V. 1, Part 1, Question 7 — *On the Infinity of God*, Cosimo Inc., New York, 2007, Art. 3, p. 32.
2. Thomas R.N. and Athay R.G. Physics of the Solar Chromosphere: Monographs and Texts in Physics and Astronomy - Vol. VI. Interscience Publishers, Inc., New York, N.Y., 1961.
3. Bray R.J. and Loughhead R.E. The Solar Chromosphere. Chapman and Hall Ltd., London, U.K., 1974.
4. Heinzel P. Understanding the solar chromosphere. In: *"Exploring the Solar System and the Universe"*, (V. Mioc, C. Dumitrache, N.A. Popescu, Eds.), American Institute of Physics, 2008, 238–244.
5. Athay R.G. The solar chromosphere and corona: Quiet Sun. D. Reidel Publishing Company, Dordrecht, Holland, 1976.
6. Secchi A. Le Soleil (2nd Edition, Part II). Guathier-Villars, Paris, 1877.
7. Janssen J. Indications de quelques-uns des résultats obténus à Guntoor, pendant l'éclipse du mois d'aôut dernier. *Compte Rendus*, 1868, v. 67, 838–39.
8. Lockyer J.N. Notice of an observation of the spectrum of a solar prominence. *Proc. Roy. Soc. London*, 1868, v. 17, 91–92.
9. Evershed J. Wave-length determinations and general results obtained from a detailed examination of spectra photographed at the solar eclipse of January 22, 1898. *Phil. Trans. Roy. Soc. London*, 1901, v. 197, 381–413.
10. Evershed J. Preliminary report of the expedition to the south limit of totality to obtain photographs of the flash spectrum in high solar latitudes. *Proc. Roy. Soc. London*, 1900, v. 67, 370–385.
11. Menzel D.H. A Study of the Solar Chromosphere. *Publications of the Lick Observatory*, University of California Press, Berkeley, CA, v. 17, 1931.
12. Menzel D.H. and Cillié G.G. Hydrogen emission in the chromosphere. *Astrophys. J.*, 1937, v. 85, 88–106.
13. Hale G.E. and Adams W.S. Photography of the "flash" spectrum without an eclipse. *Astrophys. J.*, 1909, v. 30, 222–230.
14. Adams W.S. and Burwell C.G. The flash spectrum without an eclipse region λ4800–λ6600. *Astrophys. J.*, 1915, v. 41, 116–146.
15. Wood B.E., Karovska M., Cook J.W., Brueckner G.E., Howard R.A., Korendyke C.M. and Socker D.G. Search for brightness variations in FeXIV coronagraph observations of the quiescent solar corona. *Astrophys. J.*, 1998, v. 505, 432–442.
16. Habbal S.R., Druckmüller M., Morgan H., Daw A., Johnson J., Ding A., Arndt M., Esser R., Rušin V. and Scholl I. Mapping the distribution of electron temperature and Fe charge states in the corona with total solar eclipse observations. *Astrophys. J.*, 2010, v. 708, 1650–1662.
17. Habbal S.R., Druckmüller M., Morgan H., Scholl I., Rušin V., Daw A., Johnson J. and Arndt M. Total solar eclipse observations of hot prominence shrouds. *Astrophys. J.*, 2010, v. 719, 1362–1369.
18. Habbal S.R., Morgan H. and Druckmüller M. A new view of coronal structures: Implications for the source and acceleration of the solar wind – First Asia-Pacific Solar Physics Meeting. *ASI Conf. Ser.*, 2011, v. 2, 259–269.
19. Baum W.A, Johnson F.S., Oberly J.J., Rockwood C.C., Strain C.V. and Tousey R. Solar ultraviolet spectrum to 88 km. *Phys. Rev.*, 1946, v. 70, 781–782.
20. Tousey R. The extreme ultraviolet spectrum of the Sun. *Space Sci. Reviews*, 1963, v. 2, 3-69.
21. Zirin H. and Dietz R.D. The structure of the solar chromosphere I. A picture based on extreme ultraviolet, millimeter, and λ 10830 data. *Astrophys. J.*, 1963, v. 138, 664–679.
22. Pottasch S.R. On the interpretation of the solar ultraviolet emission line spectrum. *Space Sci. Reviews*, 1964, v. 3, 816–855.
23. Pottasch S.R. On the iron lines observed in the solar ultraviolet spectrum. *Bull. Astr. Inst. Netherlands*, 1966, v.18, 237-246.
24. Goldberg L. Ultraviolet and X-Rays from the Sun. *Ann. Reviews Astron. Astrophys.*, 1967, v. 5, 279–324.
25. Doscheck G.A., Meekins J.F., Kreplin R.W., Chubb T.A., and Friedman H. Iron-line emission during solar flares. *Astrophys. J.*, 1971, v. 170, 573–586.
26. Dwivedi B.N. EUV spectroscopy as a plasma diagnostic. *Space Sci. Reviews*, 1994, v. 65, 289–316.
27. Feldman U. and Widing K.G. Elemental abundances in the solar upper atmosphere derived from spectroscopic means. *Space Sci. Reviews*, 2003, v. 107, 665-720.
28. Phillips K.J.H., Feldman U. and Landi E. Ultraviolet and X-Ray Spectroscopy of the Solar Atmosphere: Cambridge Astrophysics Series - Vol. 44, Cambridge University Press, Cambridge, U.K., 2008.
29. Zirin H. The mystery of the chromosphere. *Solar Phys.*, 1996, v. 169, 313–326.
30. Robitaille P.M. Liquid Metallic Hydrogen: A Building Block for the Liquid Sun. *Progr. Phys.*, 2011, v. 3, 60–74.
31. Robitaille J.C. and Robitaille P.M. Liquid Metallic Hydrogen III. Intercalation and Lattice Exclusion Versus Gravitational Settling and Their Consequences Relative to Internal Structure, Surface Activity, and Solar Winds in the Sun. *Progr. Phys.*, 2013, v. 2, 87-97.
32. Robitaille P.M. Commentary on the liquid metallic hydrogen model of the Sun. Insight relative to coronal rain and splashdown events. *Progr. Phys.*, 2013, v. 2, L10–L11.

33. Robitaille P.M., Abduljalil A.M., Kangarlu A., Zhang X., Yu Y., Burgess R., Bair S., Noa P., Yang L., Zhu H., Palmer B., Jiang Z., Chakeres D.M. and Spigos D. Human magnetic resonance imaging at 8 T. *NMR Biomed.*, 1998, v. 11, no. 6, 263–265.

LETTERS TO PROGRESS IN PHYSICS

The Liquid Metallic Hydrogen Model of the Sun and the Solar Atmosphere I. Continuous Emission and Condensed Matter Within the Chromosphere

Pierre-Marie Robitaille

Department of Radiology, The Ohio State University, 395 W. 12th Ave, Columbus, Ohio 43210, USA.
robitaille.1@osu.edu

The continuous spectrum of the solar photosphere stands as the paramount observation with regard to the condensed nature of the solar body. Studies relative to Kirchhoff's law of thermal emission (e.g. Robitaille P.-M. Kirchhoff's law of thermal emission: 150 years. *Progr. Phys.*, 2009, v. 4, 3–13.) and a detailed analysis of the stellar opacity problem (Robitaille P.M. Stellar opacity: The Achilles' heel of the gaseous Sun. *Progr. Phys.*, 2011, v. 3, 93–99) have revealed that gaseous models remain unable to properly account for the generation of this spectrum. Therefore, it can be stated with certainty that the photosphere is comprised of condensed matter. Beyond the solar surface, the chromospheric layer of the Sun also generates a weak continuous spectrum in the visible region. This emission exposes the presence of material in the condensed state. As a result, above the level of the photosphere, matter exists in both gaseous and condensed forms, much like within the atmosphere of the Earth. The continuous visible spectrum associated with the chromosphere provides the twenty-sixth line of evidence that the Sun is condensed matter.

> *In order to explain the occurrence of the dark lines in the solar spectrum, we must assume that the solar atmosphere incloses a luminous nucleus, producing a continuous spectrum, the brightness of which exceeds a certain limit. The most probable supposition which can be made respecting the Sun's constitution is, that it consists of a solid or liquid nucleus, heated to a temperature of the brightest whiteness, surrounded by an atmosphere of somewhat lower temperature.*
>
> Gustav Robert Kirchhoff, 1862 [1]

When Gustav Kirchhoff was contemplating the origin of the solar spectrum [1], he was probably unaware that structures beyond the photosphere also had the ability to emit continuous spectra. Still, he understood that continuous thermal emission was a property of the condensed state [1]. Gases emit in bands [2] and even compressed gases cannot produce the required thermal spectrum, outside the confines of an enclosure and in the absence of a perfect absorber [3, 4].

Despite these physical realities, over the course of the past 150 years, scientists have moved away from Kirchhoff's realization that the solar surface must be comprised of condensed matter. Instead, gaseous solar models were adopted (e.g. [5, 6]). Sadly, Kirchhoff himself enabled this misstep through his erroneous formulation of the law of thermal emission (see [3,4] and references therein). Discounting problems with the law of emission [3, 4], it can be said that the gaseous models have been based on a false premise: that the thermal spectrum of the Sun could be generated using a vast combination of non-thermal processes [7]. The solar opacity problem [7] reflects the fact that the gaseous models can never properly account for the thermal spectrum. The generation of a continuous solar spectrum has become an insurmountable hurdle for these models [7]. Though gaseous opacity calculations have been used in an attempt to account for the Sun's emission, such calculations are of no value in mirroring the simple graphitic spectrum on Earth, which the solar spectrum strongly emulates. Therefore, gaseous opacity calculations cannot have any lasting merit in generating the continuous spectrum of the Sun [7]. The emission of a thermal spectrum requires an underlying thermal mechanism, not a large sum of non-thermal processes [7]. Condensed matter is required, as illustrated by earthly black bodies (see [3,4] and references therein). In fact, the continuous spectrum of the Sun acts as the most important line of evidence that the Sun is condensed matter (see [8–14] and references therein). This was recognized long ago by Gustav Kirchhoff: gases cannot properly account for the solar spectrum [1].

The presence of continuous thermal emission by the photosphere is complemented in the outer atmosphere of the Sun. The chromosphere also supports weak continuous emission. Hence, an additional line of evidence that the Sun is comprised of condensed matter can be harvested by extending Kirchhoff's insight to the solar atmosphere, above the photospheric surface.

The weak continuous spectrum of the chromosphere [15–18] has drawn the attention of solar observers for over 100 years [19–22]. The great astronomer, Donald Howard Menzel [23], commented as follows on its nature: *"...we assumed that the distribution in the continuous chromospheric spectrum is the same as that of a black body at 5700^o, and*

that the continuous spectrum from the extreme edge is that of a black body at 4700°. There is evidence in favor of a lower temperature at the extreme limb in the observations by Abbot, Fowle, and Aldrich of the darkening towards the limb of the Sun" [22]. From early days, the continuous chromospheric spectrum was known to vary in temperature with height [24–27]. Consequently, solar observers rapidly introduced temperature variations with increasing height into their atmospheric models (e.g. [17, p. 187-213]; [18, p. 271–352]; [24–31]).

At the same time, problems remained surrounding the formation of the weak continuous chromospheric spectrum. This layer of the Sun, in the context of the modern gaseous models, had an average density of only $\sim 10^{-12}$ g/cm^3 [32, p. 32]. In fact, as one proceeds out from the photosphere to the top of the chromosphere, the density was hypothesized to be changing from $\sim 10^{-7}$ g/cm^3 to $\sim 10^{-15}$ g/cm^3, respectively [33]. It was known that in the chromosphere *"...the intensity of the emitted radiation is several tens of thousand times less than that of the photosphere"* [32, p. 32]. As a result, since the gas models were reducing photospheric densities to the levels of laboratory vacuums, the chromospheric densities had to be even lower.

In order to explain the continuous chromospheric spectrum, theoretical approaches (e.g. [16, 25–27]) exactly paralleled the methods applied for treating the emission from the photosphere (see [7] for a complete discussion). In early contributions, attention focused on neutral H, H^-, Rayleigh scattering, and electron scattering (see [17, p. 151–157] and [26, 27]). This was precisely because, devoid of condensed matter, no other mechanism could be invoked. A continuous spectrum, from which Menzel had extracted black body temperatures [22], was being explained using processes unrelated to any experimental production of a thermal spectrum on Earth [7]. Such approaches remain in use, but have already been dismissed relative to explaining the occurrence of continuous spectra [7].

Conversely, the position is now adopted that the presence of a continuous spectrum in the visible range within the chromosphere [15–18] represents a direct manifestation of condensed matter in this region of the solar atmosphere. The proper means of explaining continuous emission in the visible region of the electromagnetic spectrum, especially when it can be hypothesized to hold a thermal lineshape [22], will always remain linked to the presence of condensed matter [7].

The chromosphere corresponds to a region of the Sun where hydrogen atoms are re-entering the condensed state, prior to their recombination with photospheric material.

However, unlike the liquid metallic hydrogen advanced to be present in the solar body [8–14], chromospheric condensed matter appears to lack metallic properties. Chromospheric material, though in the condensed state, might therefore be substantially different than photospheric material. Nonetheless, though the continuous spectrum of the chromosphere remains weak, it demonstrates the presence of condensed matter within a gaseous matrix, much like drops of water can exist within the gaseous atmosphere of the Earth. In this regard, the intensity of the chromospheric emission spectrum can provide some sense of material densities in this layer. The presence of a continuous visible thermal spectrum in the chromosphere thereby constitutes the twenty-sixth line of evidence (and the sixth Planckian proof [34]) that the Sun is comprised of condensed matter (see [8–14] and references therein for the others).

Dedication

This work is dedicated to Marge Marrone, for her friendship and example in leading a joyous life.

Submitted on: March 17, 2013 / Accepted on: March 20, 2013
First published online on: May 13, 2013

References

1. Kirchhoff G. The physical constitution of the Sun. In: *Researches on the Solar Spectrum and the Spectra of the Chemical Elements*. Translated by H.E. Roscoe, Macmillan and Co., Cambridge, 1862, p. 23.
2. Robitaille P.M. The little heat engine: Heat transfer in solids, liquids, and gases. *Progr. Phys.*, 2007, v. 4, 25–33.
3. Robitaille P.M. Blackbody radiation and the carbon particle. *Progr. Phys.*, 2008, v. 3, 36–55.
4. Robitaille P.M. Kirchhoff's law of thermal emission: 150 years. *Progr. Phys.*, 2009, v. 4, 3–13.
5. Bahcall J.N. and Pinsonneault M.H. Standard solar models, with and without helium diffusion, and the solar neutrino problem. *Rev. Mod. Phys.*, 1992, v. 64, no.4, 885–926.
6. Bachall J.N., Pinsonneault M.H. and Wasserburg G.J. Solar models with helium and heavy-element diffusion. *Rev. Mod. Phys.*, 1995, v. 67, no. 4, 781–808.
7. Robitaille P.M. Stellar opacity: The Achilles heel of the gaseous Sun. *Progr. Phys.*, 2011, v. 3, 93–99.
8. Robitaille P.M. A high temperature liquid plasma model of the Sun. *Progr. Phys.*, 2007, v. 1, 70–81 (also in arXiv: astro-ph/0410075).
9. Robitaille P.M. Liquid metallic hydrogen: A building block for the liquid Sun. *Progr. Phys*., 2011, v. 3, 60–74.
10. Robitaille P.M. Liquid metallic hydrogen II: A critical assessment of current and primordial helium levels in Sun. *Progr. Phys*., 2013, v. 2, 35–47.
11. Robitaille J.C. and Robitaille P.M. Liquid metallic hydrogen III. Intercalation and lattice exclusion versus gravitational settling and their consequences relative to internal structure, surface activity, and solar winds in the Sun. *Progr. Phys*., 2013, v. 2, 87–97.
12. Robitaille P.M. Commentary on the liquid metallic hydrogen model of the Sun: Insight relative to coronal holes, sunspots, and solar activity. *Progr. Phys*., 2013, v. 2, L7–L9.
13. Robitaille P.M. Commentary on the liquid metallic hydrogen model of the Sun II. Insight relative to coronal rain and splashdown events. *Progr. Phys*., 2013, v. 2, L10–L11.
14. Robitaille P.M. Commentary on the liquid metallic hydrogen model of the Sun III. Insight into solar lithium abundances. *Progr. Phys*., 2013, v. 2, L12–L13.
15. Menzel D.H. A Study of the Solar Chromosphere. *Publications of the Lick Observatory*, University of California Press, Berkeley, CA, v. 17, 1931.

16. Thomas R.N. and Athay R.G. Physics of the Solar Chromosphere. Interscience Publishers, New York, N.Y., 1961.
17. Bray R.J. and Loughhead R.E. The Solar Chromosphere, Chapman and Hall, London, U.K., 1974.
18. Athay R.G. The Solar Chromosphere and Corona: Quiet Sun – Astrophysics and Space Science Library – v. 53. D. Reidel Publishing Company, Dordrecht, Holland, 1976.
19. Evershed J. Wave-length determinations and general results obtained from a detailed examination of spectra photographed at the solar eclipse of January 22, 1898. *Phil. Trans. Roy. Soc. London*, 1901, v. 197, 381–413.
20. Evershed J. Preliminary report of the expedition to the south limit of totality to obtain photographs of the flash spectrum in high solar latitudes. *Proc. Roy. Soc. London*, 1900, v. 67, 370–385.
21. Grotian W. Über die intensitätsverteilung des kontinuierlichen spektrums der inneren korona. *Zeitschrift fur Astrophysik*, 1931, v. 3, 199–226.
22. Menzel D.H. and Cillié G.G. Hydrogen emission in the chromosphere. *Astrophys. J.*, 1937, v. 85, 88–106.
23. Goldberg L. and Aller L.H. Donald Howard Menzel 1901–1976. In: *Biographical Memoires*, The National Academy of Sciences USA, 1991 (accessed online on 2/11/2013).
24. Athay R.G., Billings D.E., Evans J.W. and Roberts W.O. Emission in hydrogen Balmer lines and continuum in flash spectrum of 1952 total solar eclipse at Karthoum, Sudan. *Astrophys. J.*, 1954, v. 120, 94–111.
25. Athay R.G., Menzel D.H., Pecker J.C., and Thomas R.N. The thermodynamic state of the outer solar atmosphere V. A model of the chromosphere from the continuous emission. *Astrophys. J. Suppl. Ser.*, 1955, v. 1, 505–519.
26. Hiei E. Continuous spectrum in the chromosphere. *Publ. Astron. Soc. Japan*, 1963, v. 15, 277–300.
27. Weart S.R. and Faller J.E. Photoelectric eclipse observation of the continuum at the extreme solar limb. *Astrophys. J.*, 1969, v. 157, 887–901.
28. Gingerich O. and de Jager C. The Bilderberg model of the photosphere and low chromosphere. *Solar Phys.*, 1968, v. 3, 5–25.
29. Gingerich O., Noyes R.W., Kalkofen W. and Cuny Y. The Harvard-Smithsonian reference atmosphere. *Solar Phys.*, 1971, v. 18, 347–365.
30. Athay R.G. Boundary conditions on model solar chromospheres. *Solar Phys.*, 1969, v. 9, 51-55.
31. Fontenla J.M., Balasubramaniam K.S. and Harder J. Semiemperical models of the solar atmosphere II. The quiet-Sun low chromosphere at moderate resolution. *Astrophys. J.*, 2007, v. 667, 1243–1257.
32. Bhatnagar A. Instrumentation and observational techniques in solar astronomy. In: *Lectures on Solar Physics* (H.M. Antia, A. Bhatnagar and R. Ulmschneider, Eds.), Springer, Berlin, 2003, p. 27–79.
33. Ulmschneider P. The physics of the chromospheres and coronae. In: *Lectures on Solar Physics* (H.M. Antia, A. Bhatnagar and R. Ulmschneider, Eds.), Springer, Berlin, 2003, p. 232–280.
34. Robitaille P. Magnetic fields and directional spectral emissivity in sunspots and faculae: Complimentary evidence of metallic behavior on the surface of the Sun. *Progr. Phys.*, 2013, v. 1, 19–24.

LETTERS TO PROGRESS IN PHYSICS

The Liquid Metallic Hydrogen Model of the Sun and the Solar Atmosphere II. Continuous Emission and Condensed Matter Within the Corona

Pierre-Marie Robitaille

Department of Radiology, The Ohio State University, 395 W. 12th Ave, Columbus, Ohio 43210, USA.
robitaille.1@osu.edu

The K-corona, a significant portion of the solar atmosphere, displays a continuous spectrum which closely parallels photospheric emission, though without the presence of overlying Fraunhofer lines. The E-corona exists in the same region and is characterized by weak emission lines from highly ionized atoms. For instance, the famous green emission line from coronium (FeXIV) is part of the E-corona. The F-corona exists beyond the K/E-corona and, like the photospheric spectrum, is characterized by Fraunhofer lines. The F-corona represents photospheric light scattered by dust particles in the interplanetary medium. Within the gaseous models of the Sun, the K-corona is viewed as photospheric radiation which has been scattered by relativistic electrons. This scattering is thought to broaden the Fraunhofer lines of the solar spectrum such that they can no longer be detected in the K-corona. Thus, the gaseous models of the Sun account for the appearance of the K-corona by distorting photospheric light, since they are unable to have recourse to condensed matter to directly produce such radiation. Conversely, it is now advanced that the continuous emission of the K-corona and associated emission lines from the E-corona must be interpreted as manifestations of the same phenomenon: condensed matter exists in the corona. It is well-known that the Sun expels large amounts of material from its surface in the form of flares and coronal mass ejections. Given a liquid metallic hydrogen model of the Sun, it is logical to assume that such matter, which exists in the condensed state on the solar surface, continues to manifest its nature once expelled into the corona. Therefore, the continuous spectrum of the K-corona provides the twenty-seventh line of evidence that the Sun is composed of condensed matter.

> *In order to explain the occurrence of the dark lines in the solar spectrum, we must assume that the solar atmosphere incloses a luminous nucleus, producing a continuous spectrum, the brightness of which exceeds a certain limit. The most probable supposition which can be made respecting the Sun's constitution is, that it consists of a solid or liquid nucleus, heated to a temperature of the brightest whiteness, surrounded by an atmosphere of somewhat lower temperature.*
>
> Gustav Robert Kirchhoff, 1862 [1]

Providence has made of the pastoral State of Iowa one of the most important locations in the history of solar physics. From primitive observatories in Des Moines and Burlington respectively, William Harkness and Charles Young monitored the total eclipse of August 7, 1869 [2, 3], an event which still has the power to redefine our understanding of the corona.

From the heart of Iowa, William Harkness *"obtained a coronal spectrum that was continuous except for a single bright green line, later known as the coronal line K1474"* on the Kirchhoff scale [3, p. 199]. Harkness concluded that the corona was *"a highly rarefied self-luminous atmosphere surrounding the Sun, and, perhaps, principally composed of the incandescent vapor of iron"* [3, p. 199]. Eventually, John Evershed provided additional photographic evidence that the corona displayed a continuous spectrum without Fraunhofer lines and he established the wavelength of Harkness' K1474 coronal line at 5303.3 Å [4]. In addition, Evershed would document the presence of two other coronal spectral lines [2–4]. Today, the gaseous models of the Sun do not support the idea that the corona of the Sun is self-luminous. Rather, it is currently believed that the continuous coronal spectrum arises from the scattering of photospheric light by relativistic electrons in the outer solar atmosphere. In this work, Harkness' conclusion will be re-evaluated, with the intent of demonstrating that the K-corona is indeed self-luminous, as first postulated in 1869 [3, p. 196–205].

To begin understanding the corona, it is important to properly classify the spectra which it produces. It was the spectrum of the inner corona, or K-corona, which was measured long ago by Harkness, Young, and Evershed [2–4] and which has been the subject of several classic reports [5–10]. For nearly one hundred years, the inner coronal spectrum was known to be polarized [6, 10]. According to Bernard Lyot, this polarization did not extend beyond $\sim 6'$ from the limb, increased rapidly as observations were made towards the Sun,

and remained constant within ∼ 3′ of the solar surface [6].* Textbooks now state that the polarized K-corona can extend to distances approaching 10 solar radii (∼ 160′) [12, p. 187].

As Gustav Kirchhoff understood [1], Fraunhofer lines are produced when light is absorbed by gaseous atoms located above the level of the solar surface where the continuous photospheric spectrum is emitted. The absorption of light by these atoms superimposes dark lines onto the thermal spectrum of the Sun. As a result, the photospheric spectrum is always characterized by the presence of Fraunhofer lines. Conversely, these dark lines are absent in the continuous spectrum of the inner corona [13–15]. This was certainly the finding which convinced William Harkness that the corona was self-luminous [3]. For if the inner corona was simply scattering light produced by the photosphere, the Fraunhofer lines should be visible. This is the case in the outer corona, or F-corona, where photospheric light is being scattered by dust particles contained in the interplanetary medium [13, p. 33].

Within the gaseous models of the Sun, the absence of Fraunhofer lines in the K-corona is explained by scattering photospheric light with high energy electrons (see e.g. [13, p. 33] and [16]. The corona in these models has no means of directly generating a continuous spectrum. As a result, gaseous models must assume that the continuous component of coronal emission originates at the level of the photosphere. Coronal electrons must then be used to broaden the Fraunhofer lines, making them disappear from the spectrum monitored in the K-corona [13–16].

Oddly, while the gaseous models invoke electron based scattering of Fraunhofer lines, causing them to disappear in the K-corona, scattering by dust particles preserves the lines of the F-corona. The situation is further complicated because the K-corona is in the same physical space as the E-corona, which is producing emission lines, including the coronal line at 5303.3 Å [13–16]. The Fraunhofer lines are being broadened by electrons in the K-corona, but emission lines from the same region of the solar atmosphere, namely in the E-corona, remain visible and sharp. Presumably, this occurs because only a small fraction of the photospheric light is being scattered. By analogy, only a small fraction of the E-corona should be scattered. Hence, it would not be expected that the emission lines from the E-corona would be affected in a noticeable manner.

The corona is so tenuous, its emission is ∼1–100 million times less intense than that of the photosphere [13, 15]. Still, the continuous nature of its emission, and the absence of Fraunhofer lines in the inner corona has been well documented [2–10].

Speaking of the continuous coronal spectrum, Athay et al. would comment that *"It is well known (Grotrian 1931; Allen 1946) that the coronal continuum is essentially a reproduction of the photospheric continuum and does not change color with height"* [9]. Yet, Grotrian's [5] and Ludendorff's discovery (see [9]) that *"the color of the corona is the same as that of the Sun"* was not completely supported by Allen [8]. In fact, Athay [9] was misquoting Allen [8]. The latter actually found that *"microphotograms for solar distances varying from R=1.2s to R=2.6s show that the coronal radiation reddens slightly as the distance from the Sun is increased"* [8]. Allen's measurements had extended farther above the photosphere than those of Crotrian and Ludendorff, helping to explain why his predecessors had not reported reddening [8, p. 140].

Reddening of the continuous spectrum implied that the corona was cooling when one moved away from the solar surface, as would be expected. The presence of emission lines from highly ionized atoms in the E-corona appeared to be making the opposite point, the corona seemed to be much warmer than the photosphere. This issue will be addressed in detail in a separate treatment [17]. For the time being, suffice it to emphasize that the K-corona possesses a continuous spectrum which appears to be blackbody in nature and which reddens slightly with distance from the solar surface.

In the end, the simplest means of accounting for the continuous emission observed in the K-corona, the absence of overlying Fraunhofer lines, and the presence of sharp emission lines in this same region of the solar atmosphere, is to invoke a condensed matter model of the Sun [18–20]. In 1869, William Harkness had concluded that the corona was self-luminous, precisely as expected should this layer possess condensed matter.

In this regard, when the Sun is active, it is known to expel enormous amounts of material into its corona in the form of flares and coronal mass ejections. Within the liquid metallic hydrogen model of the Sun [18–20], the presence of condensed matter within the corona and the existence of an associated continuous spectrum presents little difficulty, as metallic hydrogen has already been hypothesized to be metastable (see [17] for a detailed discussion). As a result, once condensed metallic hydrogen has been produced in the solar interior, it is expected that it could retain its condensed state under the lower pressures in the corona.

The presence of condensed matter in the K-corona immediately accounts for the existence of a continuous spectrum from this region of the solar atmosphere.

At the same time, the Fraunhofer lines are not visible because insufficient levels of gaseous atoms are present in the K-corona to significantly absorb coronal radiation. Therefore, scattering by relativistic electrons does not need to be invoked to account for the presence of a continuous spectrum in the K-corona devoid of Fraunhofer lines. Conversely, the F-corona is indeed produced by the scattering of photospheric light by dust particles in interplanetary space.

As such, the continuous spectrum of the K-corona can be said to represent the twenty-seventh line of evidence that

*When visualized from the Earth, the solar diameter corresponds to ∼ 32′ or ∼ 1920″ [11]. One arc-second, (″), corresponds to ∼700 km on the Sun [12, p. 123].

the Sun is condensed matter and the seventh Planckian proof (see [21, 22] and references therein for the others).

Dedication

Dedicated to Thomas Kerner Helgeson and Barbarra Anne Helgeson who have shown the author, and provided him with, the very best from Iowa.

Submitted on: March 19, 2013 / Accepted on: March 20, 2013
Revised on April 28, 2013
First published online on: May 13, 2013

References

1. Kirchhoff G. The physical constitution of the Sun. In: *Researches on the Solar Spectrum and the Spectra of the Chemical Elements*. Translated by H.E. Roscoe, Macmillan and Co., Cambridge, 1862, p. 23.
2. Hufbauer K. Exploring the Sun: Solar Science since Galileo. The Johns Hopkins University Press, Baltimore, 1991, p. 112–114.
3. Dick S. Sky and Ocean Joined: The U.S. Naval Observatory 1830–2000. Cambridge University Press, Cambridge, 2003, p. 196-205.
4. Evershed J. Wave-length determinations and general results obtained from a detailed examination of spectra photographed at the solar eclipse of January 22, 1898. *Phil. Trans. Roy. Soc. London*, 1901, v. 197, 381–413.
5. Grotian W. Über die intensitätsverteilung des kontinuierlichen spektrums der inneren korona. *Zeitschrift für Astrophysik*, 1931, v. 3, 199–226.
6. Lyot B. La couronne solaire étudiée en dehors des éclipses. *Comptes Rendus*, 1930, v. 191, 834–837.
7. Lyot B. A study of the solar corona and prominences without eclipses – George Darwin Lecture. *Mon. Not. Roy. Astron. Soc.*, 1939, 580-=594 (22 pages with plates).
8. Allen C.W. The spectrum of the corona at the eclipse of 1940 October 1. *Mon. Not. Roy. Astron. Soc.*, 1946, v. 106, 137–150.
9. Athay R.G., Billings D.E., Evans J.W. and Roberts W.O. Emission in hydrogen Balmer lines and continuum in flash spectrum of 1952 total solar eclipse at Karthoum, Sudan. *Astrophys. J.*, 1954, v. 120, 94–111.
10. Minnaert M. On the continuous spectrum of the corona and its polarisation. *Zeitschrift für Astrophysik*, 1930, v. 1, 209–236.
11. Robitaille P.M. Commentary on the radius of the Sun. Optical illusion or manifestation of a real surface? *Progr. Phys.*, 2013, v. 2, L5-L6.
12. Zirin H. The Solar Atmosphere. Blaisdell Publishing Company, Waltham, MA, 1966.
13. Bhatnagar A. Instrumentation and observational techniques in solar astronomy. In: *Lectures on Solar Physics* (H.M. Antia, A. Bhatnagar and R. Ulmschneider, Eds.), Springer, Berlin, 2003, p. 27–79.
14. Ulmmschneider P. The physics of the chromosphere and corona. In: *Lectures on Solar Physics* (H.M. Antia, A. Bhatnagar and R. Ulmschneider, Eds.), Springer, Berlin, 2003, p. 232–280.
15. Dwivedi B.N. The solar corona. In: *Lectures on Solar Physics* (H.M. Antia, A. Bhatnagar and R. Ulmschneider, Eds.), Springer, Berlin, 2003, p. 281–298.
16. van de Hulst H.C. The electron density of the corona. *Bull. Astronom. Inst. Netherlands*, 1950, v. 11, no. 410., 135–149.
17. Robitaille P.M. The liquid metallic hydrogen model of the Sun and the solar atmosphere V. On the nature of the corona. *Progr. Phys.*, 2013, v. 3, L21–L24.
18. Robitaille P.M. Liquid metallic hydrogen: A building block for the liquid Sun. *Progr. Phys.*, 2011, v. 3, 60–74.
19. Robitaille P.M. Liquid metallic hydrogen II: A critical assessment of current and primordial helium levels in Sun. *Progr. Phys.*, 2013, v. 2, 35–47.
20. Robitaille J.C. and Robitaille P.M. Liquid metallic hydrogen III. Intercalation and lattice exclusion versus gravitational settling and their consequences relative to internal structure, surface activity, and solar winds in the Sun. *Progr. Phys.*, 2013, v. 2, 87–97.
21. Robitaille P.M. Magnetic fields and directional spectral emissivity in sunspots and faculae: Complimentary evidence of metallic behavior on the surface of the Sun. *Progr. Phys.*, 2013, v. 1, 19–24.
22. Robitaille P.M. The liquid metallic hydrogen model of the Sun and the solar atmosphere I. Continuous emission and condensed matter within the chromosphere. *Progr. Phys.*, 2013, v. 3, L5–L7.

LETTERS TO PROGRESS IN PHYSICS

The Liquid Metallic Hydrogen Model of the Sun and the Solar Atmosphere III. Importance of Continuous Emission Spectra from Flares, Coronal Mass Ejections, Prominences, and Other Coronal Structures

Pierre-Marie Robitaille

Department of Radiology, The Ohio State University, 395 W. 12th Ave, Columbus, Ohio 43210, USA.
robitaille.1@osu.edu

The solar corona and chromosphere are often marked by eruptive features, such as flares, prominences, loops, and coronal mass ejections, which rise above the photospheric surface. Coronal streamers and plumes can also characterize the outer atmosphere of the Sun. All of these structures, fascinating in their extent and formation, frequently emit continuous spectra and can usually be observed using white-light coronagraphs. This implies, at least in part, that they are comprised of condensed matter. The continuous spectra associated with chromospheric and coronal structures can be viewed as representing the twenty-eighth line of evidence, and the eighth Planckian proof, that the Sun is condensed matter. The existence of such objects also suggests that the density of the solar atmosphere rises to levels well in excess of current estimates put forth by the gaseous models of the Sun. In this work, the densities of planetary atmospheres are examined in order to gain insight relative to the likely densities of the solar chromosphere. Elevated densities in the solar atmosphere are also supported by coronal seismology studies, which can be viewed as constituting the twenty-ninth line of evidence that the Sun is composed of condensed matter.

> *In order to explain the occurrence of the dark lines in the solar spectrum, we must assume that the solar atmosphere incloses a luminous nucleus, producing a continuous spectrum, the brightness of which exceeds a certain limit. The most probable supposition which can be made respecting the Sun's constitution is, that it consists of a solid or liquid nucleus, heated to a temperature of the brightest whiteness, surrounded by an atmosphere of somewhat lower temperature.*
>
> Gustav Robert Kirchhoff, 1862 [1]

Observation of a white-light flare was initially reported by Richard Carrington in 1859 [2]. Though once considered rare events [3,4], the production of such emission has now become associated with many, if not all, flares [5]. It has been well-established that hard X-ray class flares ($\geq$ M5) emit white-light [3]. However, the mechanism for producing this light has remained elusive [6, 7], despite the prevalence of these objects [3–5]. Devoid of condensed matter, a gaseous model has little means to account for the generation of white-light flares. In 2010, Watanabe et al. [8] proposed that the emission generated by white-light flares was associated with electrons accelerated to half of the speed of light [9]. More than 150 years after Carrington's discovery, astrophysicists advanced a scenario through which white-light could be produced within the theoretical constraints imposed by accepting the idea of a gaseous Sun [10–14].

Beyond solar flares, many coronal structures are associated with the emission of white-light. These include prominences and coronal mass ejections [15–23], streamers [24–26], plumes [27], and loops [28–30]. Indeed, coronal structures have long been observed with white-light coronagraphs [25, 26], an instrument invented by Bernard Lyot [31, 32].

The existence of white-light in coronal structures presents a significant problem for the gaseous models of the Sun [10–14]. In these models, white-light at the photosphere is produced by a vast sum of processes (bound-bound, bound-free, free-free, and scattering) taking place within the Sun itself (see [33] for a complete review of this topic). In order to generate the thermal spectrum at the surface, this light must leave the hypothetically gaseous solar body through a photospheric layer regarded as an 'optical illusion' created by a dramatic change in solar opacity [34]. The current solution is so convoluted that it has been described by the author as the Achilles' Heel of gaseous solar models [33]. In no other instance is a simple spectroscopic line, such as the thermal spectrum of the Sun, produced by the extensive summation of vastly unrelated spectroscopic processes [33]. Furthermore, the mechanisms associated with the generation of the solar spectrum are of no value in explaining the thermal emission from graphite on Earth, material from which Planckian radiation was initially studied [33]. As a result, these approaches are not relevant in accounting for the thermal signature of the Sun [33].

The observation of white-light in coronal structures only

acts to accentuate this problem for the gaseous models. These objects are fleeting and devoid of the long time-lines (millions of years) currently required by the gaseous models to produce white-light from the center of the Sun. Moreover, these structures lack the large complement of processes summed within the gaseous models of the Sun to generate the white-light of the photosphere [33]. As a result, though some of the same mechanisms are invoked [3, 4], scientists who adhere to the gaseous models must now have recourse to additional effects: the scattering of photospheric light [16] or the acceleration of electrons to sub-relativistic velocities [8].

In the end, the simplest means of accounting for the presence of white-light, both on the photosphere and within coronal structures, is to recognize that the Sun is comprised of condensed matter [35–37]. The material found on the photosphere is being ejected into the solar atmosphere. Hence, it can be found within the corona. In fact, since photospheric metallic hydrogen has been hypothesized to be metastable (see [35] and references therein), it is reasonable that material ejected into the corona remains partially metallic in nature. In time, sparse filaments of condensed metallic hydrogen might come to constitute the framework for coronal streamers for instance, helping to explain why these objects also emit white-light. As a result, it is now advanced that the white-light emission of coronal structures constitutes the twenty-eighth line of evidence (see [35–39] and references therein for the others), and the eighth Planckian proof, that the Sun is comprised of condensed matter.*

Unlike the gaseous models of the Sun [10–14], the metallic hydrogen model [35–37] advances that the solar body has a nearly uniform density throughout which approaches $\sim$1 g/cm^3 at the level of the photosphere. Thus, the presence of condensed matter, expelled from the photosphere into the chromosphere and corona, strongly suggests that the densities in these regions are not negligible. In sharp contrast, within the context of a gaseous Sun and calculated electron densities, the coronal solar atmosphere is said to possess *"densities which are many trillions times smaller than that of the gas composing the Earth's atmosphere; in fact, coronal densities are low enough to be considered an almost perfect vacuum in laboratories"* [40, p. 284]. These statements are directly linked to the use of the gaseous equations of state [10, p. 130ff] and the belief that the solar body retains most of its mass in its core [10–12]. As a result, the question must naturally arise as to whether or not trillion fold decreases in densities, relative to the Earthly atmosphere, are reasonable for the solar corona. This is especially concerning relative to the realization that the Sun is expelling condensed matter [35–39] into its outer atmosphere.

*The Planckian proofs are all related to thermal emission in condensed matter. They do not imply that the objects which are the subject of these proof necessarily display a perfect thermal spectrum. The proofs are invoked when the spectrum is continuous and when an object's emissivity is most simply accounted for by invoking condensed matter.

To get some sense of reasonable densities for the corona, one can have recourse to the characteristic features of planetary atmospheres, with several important cautionary notes. First, the temperatures around the Sun and the inner planets are not at all comparable. Second, the molecular weight of material around the Sun might be either much smaller, or in the case of condensed hydrogen, *much larger*, than found in planetary atmospheres. Thirdly, the solar atmosphere might have substantial local density fluctuations well beyond anything observed in planetary atmospheres. This is especially relevant since condensed matter is being expelled into a partially gaseous solar atmosphere. These factors will impact the comparisons that can be extracted.

Consider the known densities of the Earth's atmosphere at sea level (1.229 kg/m^3 or 0.0012 g/cm^3 [41]) while taking into account that the Sun/Earth ratio of acceleration due to gravity is a factor of 28 [42]. The simple product of these values (ignoring temperature effects and assuming that the Sun's atmosphere is composed of particles of the same mean molecular weight as in the Earth's atmosphere (28.97 g/mole [43])), results in a density of 0.0336 g/cm^3 near the solar surface. This is well above current estimates for the solar atmosphere. In fact, the gaseous models of the Sun predict that, as one proceeds out from the photosphere to the top of the chromosphere, the density drops from $\sim 10^{-7}$ g/cm^3 to $\sim 10^{-15}$ g/cm^3, respectively [44, p. 32].

In reality, the aforementioned assumption that the average molecular weight in the lower solar atmosphere is similar to the Earth's cannot be correct. At the same time, temperature effects should substantially raise the amount of material found in the Sun's atmosphere. The Sun is known to expel matter into the corona and, if this is condensed matter, may have local densities well beyond that found in the atmosphere of the Earth at sea level. But even this simple calculation, based on the characteristics of the Earth's atmosphere, points to significant problems with current estimates of chromospheric densities, inferred from gaseous solar model [44] which it exceeds by a factor on the order of 10^5–10^{10}. Similar conclusions can be reached by considering Venus [45] or Mars [46].

Though some may dislike such comparisons, as too many variables could alter the final result, the author is not attempting to set a final density for the lower atmosphere of the Sun. The discussion rests simply in highlighting that the currently accepted solar values are well outside the bounds of reason, especially when considering that the Sun is much hotter than the inner planets and constantly expelling matter into its corona. This implies that a much higher average molecular weight for the solar atmosphere can be expected than one based on the atomic weight of hydrogen. Unlike the Sun, the inner planets do not eject much material into their atmospheres. As a result, the atmosphere of the Sun is likely to possess great local variability in its densities. This may also be true when comparing the atmosphere of the quiet Sun near the solar poles with that above the equator, as a result of coro-

nal holes above the former.

Finally, it remains highly significant that, when a comet approaches the Sun, it can result in intense shock wave propagation throughout the corona (e.g. [47]). Such behavior calls for highly elevated atmospheric densities. It is not reasonable to expect that shock waves and seismic activity could propagate within a corona whose density remains inferior to earthly vacuums. As such, seismological findings and shock wave propagation are highly supportive of the realization that the solar chromosphere and corona are much denser than currently surmised from the gaseous models of the Sun. Along these lines, it is concerning that the Sun can be studied using coronal helioseismology [48–51] which suggests a twenty-ninth line of evidence that it is comprised of condensed matter. It is not possible to conduct coronal seismological studies in an atmosphere sparser than the best laboratory vacuums. Seismology is a science which can be applied exclusively to the condensed states of matter.

Dedication

Dedicated to Gregory Gribbon, my longtime Canadian childhood friend, in thanksgiving for his faithfulness and support.

Submitted on: April 28, 2013 / Accepted on: May 2, 2013
First published online on: May 13, 2013

References

1. Kirchhoff G. The physical constitution of the Sun. In: *Researches on the Solar Spectrum and the Spectra of the Chemical Elements*. Translated by H.E. Roscoe, Macmillan and Co., Cambridge, 1862, p. 23.
2. Carrington R.C. Description of a singular appearance seen in the Sun on September 1, 1859. *Mon. Not. Roy. Astron. Soc.*, 1859, v. 20, 13–15.
3. Machado M.E., Emslie A.G., Avrett E.H. Radiative backwarming in white-light flares. *Solar Phys.*, 1989, v. 124, 303–317.
4. Hiei E. A continuous spectrum of a white-light flare. *Solar Phys.*, 1982, v. 80, 113–127.
5. Wang H. Study of white-light flares observed by Hinode. *Res. Astron. Astrophys.*, 2009, v. 9, no. 2, 127–132.
6. Priest E.R. Solar flare theory and the status of flare understanding. *High Energy Solar Physics: Anticipating HESSI (R. Ramaty and N. Mandzhavidze, Eds.), ASP Conf. Ser.*, 2000, v. 206, 13–26.
7. Kahler S.W. Solar flares and coronal mass ejections. *Annu. Rev. Astron. Astrophys.*, 1992, v. 30, 113–141.
8. Watanabe K., Krucker S., Hudson H., Shimizu T., Masuda S., and Ichimoto K. G-band and hard X-ray emissions of the 2006 December 14 flares observed by HINODE/SOT and RHESSI. *Astrophys. J.*, 2010, v. 715, 651–655.
9. Anderson J.L. Hinode discovers origin of white-light flares. NASA–Marshall Space Flight Center. Accessed online on 4/17/2013: http://www.nasa.gov/centers/marshall/news/news/releases/2010/10-052.html
10. Kippenhahn R. and Weigert A. Stellar structure and evolution. Springer-Verlag, Berlin, 1990.
11. Bahcall J.N. and Pinsonneault M.H. Standard solar models, with and without helium diffusion, and the solar neutrino problem. *Rev. Mod. Phys.*, 1992, v. 64, no.4, 885–926.
12. Bachall J.N., Pinsonneault M.H. and Wasserburg G.J. Solar models with helium and heavy-element diffusion. *Rev. Mod. Phys.*, 1995, v. 67, no. 4, 781–808.
13. Robitaille P.M. A thermodynamic history of the solar constitution – I: The journey to a gaseous Sun. *Progr. Phys.*, 2011, v. 3, 3–25.
14. Robitaille P.M. A thermodynamic history of the solar constitution – II: The theory of a gaseous Sun and Jeans' failed liquid alternative. *Progr. Phys.*, 2011, v. 3, 41–59.
15. Tandberg-Hanssen E. Solar prominences — An intriguing phenomenon. *Solar Phys.*, 2011, v. 269, 237–251.
16. Tandberg-Hanssen E. A spectroscopic study of quiescent prominences. *Astrophysica Novegica*, 1964, v. 9, no. 3, 13–32.
17. Yakovkin N.A. and Zeldina M.Yu. The prominence radiation theory. *Solar Phys.*, 1975, v. 45, 319–338.
18. Gopalswamy N. and Hanaoka Y. Coronal dimming associated with a giant prominence eruption. *Astrophys. J.*, 1998, v. 498, L179–L182.
19. Gopalswamy N., Shimojo M., Lu W., Yashiro S., Shibasaki K. and Howard R.A. Prominence eruptions and coronal mass ejections: A statistical study using microwave observations. *Astrophys. J.*, 2003, v. 586, 562–578.
20. Sheeley N.R., Walters J.H., Wang Y.M. and Howard R.A. Continuous tracking of coronal outflows: Two kinds of coronal mass ejections. *J. Geophys. Res.*, 1999, v. 104, no. A11, 24739–24767.
21. St. Cyr O.C., Howard R.A., Sheeley N.R., Plunkett S.P., Michels D.J., Paswaters S.E., Koomen M.J., Simnett G.M., Thompson B.J., Gurman J.B., Schwenn R., Webb D.F., Hildner E. and Lamy P.L. Properties of coronal mass ejections: SOHO LASCO observations from January 1996 to June 1998. *J. Geophys. Res.*, 2000, v. 105, no. A8, 18169–18185.
22. Hudson H.S. and Cliver E.W. Observing coronal mass ejections without coronagraphs. *J. Geophys. Res.*, 2001, v. 106, no. A11, 25199–25213.
23. Yashiro S., Gopalswamy N. Michaelek G, St. Cyr O.C., Plunkett S.P., Rich N.B. and Howard R.A. A catalog of white-light coronal mass ejections observed by the SOHO spacecraft. *J. Geophys. Res.*, 2004, v. 109, A07105(11 pages).
24. Wang Y.M. Nonradial coronal streamers. *Astrophys. J.*, 1996, v. 456, L119–L121.
25. Wang Y.M., Sheeley N.R., Walters J.H., Brueckner G.E., Howard R.A., Michels D.J., Lamy P.L., Schwenn R. and Simnett G.M. Origin of streamer material in the outer corona. *Astrophys. J.*, 1998, v. 498, L165–L168.
26. Vourlidas A. A review of white-light streamers at the end of cycle 23. *Proc. IAU: Solar Activity and its Magnetic Origin*, 2006, v. 233, 197–204.
27. Del Zanna G., Bromage B.J.I. and Mason H.E. Spectroscopic characteristics of polar plumes. *Astron. Astrophys.*, 2003, v. 398, 743–763.
28. Mouschovias T.Ch. and Poland A.I. Expansion and broadening of coronal loop transients: A theoretical explanation. *Astrophys. J.*, 1978, v. 220, 675–682.
29. Kjeldseth-Moe O. and Brekke P. Time variability of active region loops observed with the coronal diagnostic spectrometer (CDS) on SOHO. *Solar Phys.*, 1998, v. 182, 73–95.
30. Landi E., Miralles M.P., Curdt W. and Hara H. Physical properties of cooling plasma in quiescent active region loops. *Astrophys. J.*, 2009, v. 695, 221-237.
31. Lyot B. La couronne solaire étudiée en dehors des éclipses. *Comptes Rendus*, 1930, v. 191, 834–837.
32. Lyot B. A study of the solar corona and prominences without eclipses – George Darwin Lecture. *Mon. Not. Roy. Astron. Soc.*, 1939, 580-=594 (22 pages with plates).
33. Robitaille P.M. Stellar opacity: The Achilles' heel of the gaseous Sun. *Progr. Phys.*, 2011, v. 3, 93–99.
34. Robitaille P.M. On the Presence of a Distinct Solar Surface: A Reply to Hervé Faye. *Progr. Phys.*, 2011, v. 3, 75–78.

35. Robitaille P.M. Liquid Metallic Hydrogen: A Building Block for the Liquid Sun. *Progr. Phys.*, 2011, v. 3, 60–74.
36. Robitaille P.M. Liquid Metallic Hydrogen II: A Critical Assessment of Current and Primordial Helium Levels in Sun. *Progr. Phys.*, 2013, v. 2, 35–47.
37. Robitaille J.C. and Robitaille P.M. Liquid Metallic Hydrogen III. Intercalation and Lattice Exclusion Versus Gravitational Settling and Their Consequences Relative to Internal Structure, Surface Activity, and Solar Winds in the Sun. *Progr. Phys.*, 2013, v. 2, 87–97.
38. Robitaille P.M. Commentary on the liquid metallic hydrogen model of the Sun II: Insight relative to coronal rain and splashdown events. *Progr. Phys.*, 2013, v. 2, L10–L11.
39. Robitaille P.M. The Liquid Metallic Hydrogen Model of the Sun and the Solar Atmosphere II. Continuous Emission and Condensed Matter Within the Corona. *Progr. Phys.*, 2013, v. 3, L8–L10.
40. Dwivedi B.N. The solar corona. In: *Lectures on Solar Physics* (H.M. Antia, A. Bhatnagar and R. Ulmschneider, Eds.), Springer, Berlin, 2003, p. 281–298.
41. NASA. Air properties definitions. (accessed online on 2/13/2013) www.grc.nasa.gov/WWW/k-12/airplane/airprop.html
42. NASA. Sun/Earth Comparison. (accessed online on 2/13/2013) nssdc.gsfc.nasa.gov/planetary/factsheet/sunfact.html
43. NASA. Earth Fact Sheet. (accessed online on 2/13/2013) nssdc.gsfc.nasa.gov/planetary/factsheet/earthfact.html
44. Bhatnagar A. Instrumentation and observational techniques in solar astronomy. In: *Lectures on Solar Physics* (H.M. Antia, A. Bhatnagar and R. Ulmschneider, Eds.), Springer, Berlin, 2003, p. 27–79.
45. NASA. Venus/Earth Comparison. (accessed online on 2/13/2013) nssdc.gsfc.nasa.gov/planetary/factsheet/venusfact.html
46. NASA. Mars/Earth Comparison. (accessed online on 2/13/2013) nssdc.gsfc.nasa.gov/planetary/factsheet/marsfact.html
47. SOHO NASA/ESA [2011/05/10 20:00:00 to 2011/05/11 08:00:00 UTC]. These events have been captured in video format and displayed online: e.g. youtube.com/watch?v=igeBrSGk5FA; *Russia Today* youtube.com/watch?NR=1&v=Mat4dWpszoQ&feature=fvwp. (Accessed online on January 29, 2013: Examine beginning at 2011/05/10 2:48:00 to 2011/05/10 4:00:00).
48. Nakariakov V.M. and Verwichte E. Coronal waves and oscillations. *Living Rev. Solar Phys.*, 2005, v. 2, 3–65.
49. Roberts B. Progress in coronal seismology. *Proc. IAU: Waves and Oscillations in Solar Atmosphere*, 2007, v. 247, 3–19.
50. Zaqarashvilli T.V. and Erdélyi R. Oscillations and waves in solar spicules. *Space Sci. Rev.*, 2009, v. 149, 355–388.
51. De Moortel I. An overview of coronal seismology. *Phil. Trans. R. Soc. A*, 2005, v. 363, 2743–2760.

LETTERS TO PROGRESS IN PHYSICS

The Liquid Metallic Hydrogen Model of the Sun and the Solar Atmosphere IV. On the Nature of the Chromosphere

Pierre-Marie Robitaille

Department of Radiology, The Ohio State University, 395 W. 12th Ave, Columbus, Ohio 43210, USA.
robitaille.1@osu.edu

The chromosphere is the site of weak emission lines characterizing the flash spectrum observed for a few seconds during a total eclipse. This layer of the solar atmosphere is known to possess an opaque Hα emission and a great number of spicules, which can extend well above the photosphere. A stunning variety of hydrogen emission lines have been observed in this region. The production of these lines has provided the seventeenth line of evidence that the Sun is comprised of condensed matter (Robitaille P.M. Liquid Metallic Hydrogen II: A critical assessment of current and primordial helium levels in Sun. *Progr. Phys.*, 2013, v. 2, 35–47). Contrary to the gaseous solar models, the simplest mechanism for the production of emission lines is the evaporation of excited atoms from condensed surfaces existing within the chromosphere, as found in spicules. This is reminiscent of the chemiluminescence which occurs during the condensation of silver clusters (Konig L., Rabin I., Schultze W., and Ertl G. Chemiluminescence in the Agglomeration of Metal Clusters. *Science*, v. 274, no. 5291, 1353–1355). The process associated with spicule formation is an exothermic one, requiring the transport of energy away from the site of condensation. As atoms leave localized surfaces, their electrons can occupy any energy level and, hence, a wide variety of emission lines are produced. In this regard, it is hypothesized that the presence of hydrides on the Sun can also facilitate hydrogen condensation in the chromosphere. The associated line emission from main group and transition elements constitutes the thirtieth line of evidence that the Sun is condensed matter. Condensation processes also help to explain why spicules manifest an apparently constant temperature over their entire length. Since the corona supports magnetic field lines, the random orientations associated with spicule formation suggests that the hydrogen condensates in the chromosphere are not metallic in nature. Spicules provide a means, not to heat the corona, but rather, for condensed hydrogen to rejoin the photospheric layer of the Sun. Spicular velocities of formation are known to be essentially independent of gravitational effects and highly supportive of the hypothesis that true condensation processes are being observed. The presence of spicules brings into question established chromospheric densities and provides additional support for condensation processes in the chromosphere, the seventh line of evidence that the Sun is comprised of condensed matter.

> *In order to explain the occurrence of the dark lines in the solar spectrum, we must assume that the solar atmosphere incloses a luminous nucleus, producing a continuous spectrum, the brightness of which exceeds a certain limit. The most probable supposition which can be made respecting the Sun's constitution is, that it consists of a solid or liquid nucleus, heated to a temperature of the brightest whiteness, surrounded by an atmosphere of somewhat lower temperature.*
>
> Gustav Robert Kirchhoff, 1862 [1]

Nearly 150 years have now passed since Kirchhoff wrote about the Sun [1] and Father Angelo Secchi illustrated chromospheric spicules for the first time [2, p. 32]. Secchi viewed the chromospheric region as clearly defined on one side, like the surface of a liquid layer [2, p. 33]. Though he had concluded that the body of the Sun was gaseous, he believed that condensed matter was "suspended" within the photosphere [3]. Secchi would comment on the appearance of spicules and the outer portion of the chromosphere: *"In general, the chromosphere is poorly terminated and its external surface is garnished with fringes ... It is almost always covered with little nets terminated in a point and entirely similar to hair"*.* Secchi mentioned the tremendous variability in spicule orientation, their enormous size, and how these structures reminded him of flames present in a field wherein one burns grasses after the harvest ... *"It often happens, especially in the region of sunspots, that the chromosphere presents an aspect of a very active network whose surface, unequal and rough, seems composed of brilliant clouds analogous to our cumulus; the*

*All translations from French were accomplished by the author.

disposition of which resembles the beads of our rosary; a few of which dilate in order to form little diffuse elevations on the sides" [2, p. 31–36]. He would emphasize that *"there is thus no illusion to worry about, the phenomena that we have just exposed to the reader are not simple optical findings, but objects which really exist, faithfully represented to our eyes using instruments employed to observe them"* [2, p. 35–36].

The chromosphere is a region of intense magnetic activity, but its nature, and in particular that of its mottles and spicules [4–15], remains a mystery [16]. The low chromosphere is dominated by emission lines from neutral atoms and rare earths, but near its upper boundary strong lines from CaII and H are present [16]. Harold Zirin highlights that *"The chromosphere is the least-well understood layer of the Sun's atmosphere... Part of the problem is that it is so dynamic and transient. At this height an ill-defined magnetic field dominates the gas and determines the structure. Since we do not know the physical mechanisms, it is impossible to produce a realistic model. Since most of the models ignored much of the data, they generally contradict the observational data. Typical models ignore other constraints and just match only the XUV data; this is not enough for a unique solution. It reminds one of the discovery of the sunspot cycle. While most of the great 18th century astronomers agreed that the sunspot occurrence was random, only Schwabe, an amateur, took the trouble to track the number of sunspots, thereby discovering the 11-year cycle"* [16].

The struggle to understand the chromosphere is, in large measure, a direct result of the adherence to gaseous models of the Sun and a rejection of condensed matter [17–21]. The chromosphere is hypothesized to be only 2–3,000 km thick [7, p. 232]. Yet, chromospheric emission lines from hydrogen, calcium, and helium can extend up to 10,000 km above the solar surface [6, p. 8]. Zirin comments on the chromosphere as follows: *"Years ago the journals were filled with discussions of 'the height of the chromosphere'. It was clear that the apparent scale height of 1000 km far exceeded that in hydrostatic equilibrium. In modern times, a convenient solution has been found – denial. Although anyone can measure its height with a ruler and find it extending to 5000 km, most publications state that it becomes the corona at 2000 km above the surface. We cannot explain the great height or the erroneous models... While models say 2000 km, the data say* 5000″ [16].

Though the chromosphere contains bright floculi in the K line of Ca II, which coincide in position with Hα rosettes [6, p. 85–86], and though it is laced with bright/dark mottles and spicules, gaseous solar models [22–24] have no direct means of accounting for such structures [25].

It remains fascinating that spicule formation velocities appear to be largely independent of gravitational forces [9–15], though some efforts have been made to establish such a relationship [26]. In general, while most velocities of spicules formation seem to move at nearly uniform speeds [4, p. 61], some actually increase with elevation, rise in jerks, or stop suddenly upon reaching their maximum height [6, p. 45–60]. Spicules have been said to *"expand laterally or slit into two or more strands after being ejected"* [26]. Such behavior is strongly suggestive of a condensation process.

Spicules are often associated with magnetic phenomena in the chromosphere [4–15]. They can be represented as lying above photospheric intergranular lanes. In so doing, they seem to be experiencing lateral magnetic pressure from the material trapped within the field lines that originate in the solar surface, as displayed in Fig. 1.

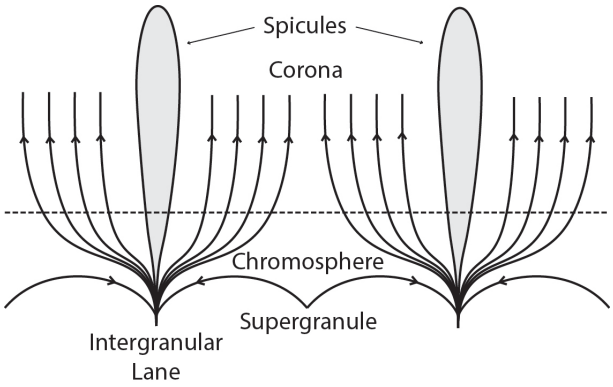

Fig. 1: Schematic representation of spicules overlying the intergranular lane on the outer boundary of a supergranule and surrounded by magnetic field lines emanating from the solar surface. This figure is an adaptation based on Fig. IV–13 in [4, p. 162].

Numerous magnetic field lines escape from the solar interior through the photospheric surface. These fields must traverse the chromospheric material. As a result, most solar observers believe that chromospheric structures are inherently magnetic [4–15]. Spicules are though to be propelling matter upwards into the corona [27] and not gathering matter, through condensation, for rejoining the photosphere.

However, given the appearance of chromospheric structures, such as rosettes and mottles, and the somewhat random orientations of spicules [4–15], it seems unlikely that these objects can be of magnetic origin. What is more probable is that, while non-metallic, chromospheric structures are being confined by charged plasmas, or metallic hydrogen [17–21], flowing in conjunction with the solar magnetic fields lines, much as illustrated in Fig. 1.

Since the gaseous models [22–24] depend on excessive temperatures in order to explain emission lines, spicules have been advanced as partly responsible for heating the corona [27]. Two forms of spicules were postulated from observations. Type I spicules can be viewed as classic spicules with lifetimes on the order of 3–7 minutes [27]. Type II spicules were believed to form rapidly, be short lived (10–150 s) and thin (<200 km), and capable of projecting material into the upper chromosphere at great velocities [27]. Type I spicules

were said to move up and down, while Type II spicules faded [27]. Type II spicules were claimed to be potentially important in heating the outer atmosphere [27]. But recently, their existence and role appears to have been soundly refuted [28].

Though chromospheric observers remain intrigued with structure [4–15, 26], they adhere to the gaseous solar models [22–24], even though gases cannot exhibit true condensation. As a result, all chromospheric and coronal structures must be viewed as gasoues plasmas of exceedingly low densities [7]. Since they are not condensed matter in the context of the gaseous models [22–24], the strange properties of spicule formation and the structures of rosettes and mottles, remain an anomaly, rather than indicators of the nature of the chromosphere.

In contrast, this work now advances that chromospheric structures represent solar material in the condensed state [17–21]. Matter in this region fluctuates between gaseous and condensed, as spicules and mottles form and dissipate [9–15, 26]. This is reminiscent of phenomena such as critical opalescence [25]. The chromosphere appears to be a site of hydrogen condensation. The mystery lies only in how this can be achieved.

In order to better understand the chromosphere, one can revisit the classic work of Donald H. Menzel published 1931, *"A Study of the Solar Chromosphere"* [29]. Within this volume, three revelations continue to make their mark. First, there is an amazing prevalence of emission lines from a wide variety of atoms within this layer of the Sun (see Table I [29, p. 18–113]). Second, the chromosphere contains an extensive group of emission lines from hydrogen (see Table 3 [29, p. 128]). Menzel lists more than twenty-three hydrogen emission lines in his Table 3 [29]. Along with Cillié, Menzel soon observed Balmer series emission up to H31, with higher states limited only by resolution [30]. Third, he outlines a hydrogen abundance in the chromosphere which is 100 times more elevated than in the Sun (see Table 20 [29, p. 281]). Menzel's chromospheric hydrogen abundance was nearly 1,000 times more elevated [29, p. 275–281] than that which had been reported by Henry Norris Russell from the Fraunhofer spectrum of the Sun itself just a few years before [31].

While eighty years have passed since *"A Study of the Solar Chromosphere"* was published [29], much remains to be understood relative to this region of the Sun. P. Heinzel writes that *"Moreover, the energy supply into these layers is largely unknown, we only know that the radiation is not the dominant source of heating. The solar chromosphere is probably the least understood part of the Sun, even compared to the solar interior on which helioseismology has focused during last decades"* [8]. Much like other physical processes in the Sun, local heating is being tentatively attributed to magnetic mechanisms. Yet, if the chromosphere remains a mystery, the cause rests on the insistence that the Sun must exist in the gaseous state. Donald Menzel reminds us, *"The province of the scientist is the untangling of mysteries, the rendering of complex things into simple ones"* [29, p. 1].

A condensed solar model provides elegant solutions to the most perplexing questions relative to the chromosphere. This is especially true relative to apparent heating, as best understood through the careful consideration of how chromospheric emission lines are produced.

Within gaseous models, line emission requires either photon absorption, or electron collision, to excite the emitting atom [4, p. 228]. This represents an attempt to explain spectra using random processes. In the condensed model [17–21] spectra are tied to the formation of chromospheric structures. Line emission becomes inherently linked to understanding the very nature of the chromosphere.

The quest for answers begins with the consideration of condensation processes in clusters, the smallest precursors to condensed matter [32–37]. Clusters can be super-stable and act as superatoms [38]. In addition, their most favorable configurations can be linked to highest electron affinity and not to the energy of the ground state [39]. Condensation processes in clusters have been known to be associated with light emission [40, 41] and are exothermic. Thus, the apparent heating of the chromosphere might best be understood by considering these reactions.

In 1996, chemiluminescence was first reported to occur during the agglomeration of silver clusters [40]. By necessity, the reactions involved took place at low temperature (~30K), but the lessons learned directly translate to other conditions. Gerhart Ertl (Nobel Prize, Chemistry, 2007) and his team highlight: *"Exothermic chemical reactions may be accompanied by chemiluminescence. In these reactions, the released energy is not adiabatically damped into the heat bath of the surrounding medium but rather is stored in an excited state of the product; decay from this excited state to the ground state is associated with light emission"* [40].

The reactions presented by Ertl [40], which are of interest relative to the chromosphere, are illustrated by the condensation of two silver fragments, resulting in an activated cluster species: $M_n + M_m \rightarrow M^*_{m+n}$. The activated cluster returns to the ground state by ejecting an excited atom: $M^*_{m+n} \rightarrow M_{m+n-1} + M^*$. Finally, the excited silver atom is able to relax to the ground state by emitting light: $M^* \rightarrow M + hv$. Consequently, since condensation processes are exothermic, they are capable of producing excited atoms which result in emission. To extend these concepts to the solar chromosphere, it is useful to consider the types of condensation reactions which might be present in this region of the Sun.

In the chromosphere, it is possible to observe spectroscopic emission lines from atomic hydrogen corresponding to the Lyman ($n_2 > 1 \rightarrow n_1 = 1$ [42]), Balmer ($n_2 > 2 \rightarrow n_1 = 2$ [30]), and Paschen series ($n_2 > 3 \rightarrow n_1 = 3$ [43]).* Lyman emission lines involve relaxation back to the ground state and

*Up to eleven separate Lyman emission lines have been recorded (n_2=7 through n_2=17 [42, p. 47]), Balmer lines at least up to n_2=31 [30], and at least nine Paschen lines (n_2=8 through n_2=16) [43]

can directly be deduced to arise from the condensation of hydrogen fragments, $H_n + H_m \rightarrow H^*_{m+n}$, relaxation of the resultant condensation product through the ejection of an excited hydrogen atom, $H^*_{m+n} \rightarrow H_{m+n-1} + H^*$, and finally the return to the ground state of the excited hydrogen atom with light emission, $H^* \rightarrow H + hv$.

In reality, it is reasonable to postulate that reactions in the chromosphere primarily involves the combination of molecular hydrogen with much larger condensed hydrogen structures or seeds, CHS, since this region of the Sun displays tangible signs of condensed matter in the form of spicules and mottles [9–15, 26].

In this case, molecular hydrogen, H_2, initially combines with these larger structures, $CHS + H_2 \rightarrow CHS–H^*_2$, resulting in mass increase, CHS–H, and the subsequent line emission from the ejected hydrogen atom, $H^* \rightarrow H + hv$. Given the extensive quantities of hydrogen in the Sun, it would be expected that numerous such reactions could take place simultaneously on any given CHS and result in the rapid appearance of spicules and mottles in the solar atmosphere [9–15, 26]. Since these reactions are *"not adiabatically dumped into the heat bath of the surrounding medium"* [40], condensation processes could result in the emission of Balmer [30] and Paschen lines [43]. In fact, the first line of the Balmer series ($n_2 = 3 \rightarrow n_1 = 2$), known as the H_α line, is responsible for the reddish hue of the chromosphere [7, p. 232].

The aforementioned reactions depend on the presence of molecular hydrogen in the chromosphere. Unfortunately, the concentrations of molecular hydrogen are extremely difficult to estimate in astrophysics, even if this species is widely considered to the most abundant molecule in the universe. The difficulty in establishing molecular hydrogen concentrations stems from the fact that all rotational-vibrational transitions from the ground electronic state of this diatom are forbidden [44]. As a result, astronomers typically use indirect methods to compute molecular hydrogen fractions in the galaxies [44–47] and sunspot umbra where the molecule is thought to be abundant [48].

Nonetheless, molecular hydrogen has been directly observed in sunspots in the extreme ultra-violet, a region of the electromagnetic spectrum were the emission lines are relatively strong [49, 50]. Furthermore, Jordan et al. [49] report a significant enhancement of the molecular hydrogen signal when chromospheric material lies over the sunspot of interest. While these signals are generally weak on the quiet Sun and in the limb [51], the emission from flares [50] and chromospheric plages [52] can be rather strong. Given the difficulty in observing molecular hydrogen in the ground state, these findings are significant and highlight that this species should be available to support condensation reactions in the chromosphere.

Therefore, it is likely that the hydrogen emission lines at the chromospheric level are related to the growth of CHS and the recapture of hydrogen from the outer solar atmosphere. Przybilla and Buttler have already simulated the linewidth of hydrogen emission lines in the chromosphere and reached the conclusion that some of the lines *"couple tightly to the continuum"* [53]. But within the context of the gaseous solar models, it is impossible to *"couple tightly to the continuum"*, as the latter merely represents an opacity change, not a physical structure [54]. It is for this reason that the emission lines of hydrogen have already been ascribed to the seventeenth line of evidence that the Sun is comprised of condensed matter [19]. Line emission can be linked to condensation, as Ertl has already elegantly demonstrated [40, 41]. Moreover, within the condensed models of the Sun, it would be natural that hydrogen emission associated with condensation would *"couple tightly to the continuum"*.

Before closing the discussion of hydrogen, it is important to digress slightly from addressing emission in order to discuss the hydrogen Fraunhofer absorption lines of the Balmer series. These lines are known to be broad and, as first reported by Unsöld [55], their relative intensities do not decrease in the manner predicted from quantum mechanical considerations. This has already been discussed by the author [19]. Therefore, Fraunhofer lines are not directly related to condensation processes. Isolated atoms, unlike diatoms, lack the ability to add protons to condensed structures, while at the same time removing heat. It is unlikely that isolated atoms can condense onto larger structures. It is more probable that they combine with one another to make a molecular species which, in turn, can condense. Hence, the broadening associated with the Balmer Fraunhofer lines can be linked to collisional processes whereby atoms are strongly interacting with the condensed matter which surrounds them, but not condensing. This represents, as previously mentioned, the sixteenth line of evidence that the Sun is comprised of condensed matter [19].

Returning to line emission, in addition to molecular hydrogen, the chromosphere may well possess other species which can facilitate the condensation of hydrogen atoms. Indeed, many hydrides have been identified either on the solar disk itself or within sunspots [56, 57], including CaH, MgH, CH, OH, H_2O, NH, SH, SiH, AlH, CoH, CuH, and NiH. The presence of CaH and MgH in the Sun have been known since the beginning of the 20th century [58]. In the laboratory hydrides from the main group elements (Li, Na, K, Rb, Cs, Be, Mg, Ca, St, Ba, B, Al, Ga, In, Tl, C, Si, Ge, Sn, N, P, As, Sb, Bi, O, S, Se, Te, F, Cl, Br, and I) and many of the transition metals (including amongst others V, Fe, Co, Ni, Cu, Ag, Zn, and Cd) are readily synthesized [59]. Hydrogen appears to have a great disposition to form hydrides of all kinds and this is an important realization relative to understanding the lower solar atmosphere.

Interestingly, the emission lines from CaII and MgII are particularly important in the chromosphere (e.g. [4, p. 361–369]). The second ionization state is singly charged (Ca^+ and Mg^+). But, the inert gas structure of these ions would demand

a doubly charged species, i.e. Ca^{+2} and Mg^{+2}. As such, why is it that the most important ions of calcium and magnesium on the Sun are singly charged? The answer is likely to rest with their role in making hydrogen available for condensation.

Consider the reactions for calcium. It should be possible for CaH and a condensed hydrogen structure to create an activated complex, CHS + CaH → CHS–HCa*. This would then be followed by an exothermic step involving the expulsion of an activated CaII ion, CHS–HCa → CHS–H + Ca^{+*}, followed by the line emission from CaII*, $Ca^{+*} \rightarrow Ca^+ + hv$.

An identical scenario could be advanced for all the monohydrides, resulting in the observed line emission from their associated cations. Indeed, chromospheric emission lines, involving cations in modest oxidation states, are likely to be generated following a very similar mechanism. Some atoms, like oxygen or iron, may well exist as dihydride or higher complexes of hydrogen. They should participate similarly in condensation reactions, bringing in the process one or more hydrogen atoms to the site of condensation. The metal hydrides thereby would constitute important building blocks in the resynthesis of condensed forms of hydrogen.

When molecular hydrogen delivers a single proton to the condensation reactions, it is also delivering a single electron, if a neutral hydrogen atom subsequently emits. The same can be said for all hydrides wherein neutral atoms are ejected from the condensate to then produce emission lines. Atoms like oxygen have higher ionization potentials that the alkali, alkaline, or transition metals and may well prefer to hold on to their electrons. Emission lines from neutral oxygen are well know to be present during spicule formation [60].

Conversely, a species like CaH is delivering two electrons when generating CaII, as the negative hydrogen ion is being released. This suggests that condensed hydrogen structures, CHS, in the chromosphere might have reasonable electron affinities, though perhaps slightly less than that of oxygen in the lower chromosphere.

Importantly, the delivery of hydrogen to condensed hydrogen structures will involve potentially strong interactions between the carrier atoms (H, Ca, Mg, etc.) and the condensate surface. This would be expected to result in substantial line broadening of the ejected excited species. In support of such an idea, CaII and MgII spicule lines are known to be broad, and the H_α emission lines also display increased linewidths (see e.g. [60–62]). Such findings suggest tight coupling of these atoms to the condensate prior to ejection. Conversely, spicule emission linewidths from the H_β, H_γ, H_ϵ emission line, the D3 line from He, and the line from neutral oxygen are all sharp [60] in spicules, suggesting weaker coupling in those cases.

Contrary to gaseous models of the Sun which have ascribed no reasonable function to the chromosphere, the liquid metallic hydrogen framework [17–21] appears to provide a sound purpose for this layer. A condensed Sun does not permit hydrogen to simply escape, without recovery, into extrasolar space. Rather, molecular hydrogen and hydrides are likely to be participating in the continued recondensation of hydrogen within the chromosphere generating the observed emission lines. The resulting material appears to be nonmetallic since spicules can display orientations which are not coupled to the magnetic field lines of the Sun [9–15]. This material may then rejoin the photosphere and travel into the solar interior, perhaps using intergranular lanes [63]. Once in the interior of the Sun, pressure would facilitate the resynthesis of metallic hydrogen.

In summary, for the first time, it is advanced that complex condensation reactions take place in the chromosphere. These result in line-emission and provides a novel way to explain both spectra and structures on the Sun. The chromosphere appears to be rich in atomic and molecular hydrogen. Furthermore, a wide array of hydride based reactions seem to occur within the chromosphere and these provide a powerful incentive to further the understanding of condensation and hydride chemistry on Earth. In this respect, the presence of metal hydrides [56–58] and the line emission of main group and transition elements in the chromosphere constitutes the thirtieth line of evidence that the Sun is comprised of condensed matter.

Acknowledgment

Luc Robitaille is acknowledged for the figure preparation.

Dedication

Dedicated to the poor, who sleep, nearly forgotten, under the light of the Southern Cross.

Submitted on: May 1, 2013 / Accepted on: May 2, 2013
First published in online on: May 13, 2013

References

1. Kirchhoff G. The physical constitution of the Sun. In: *Researches on the Solar Spectrum and the Spectra of the Chemical Elements*. Translated by H.E. Roscoe, Macmillan and Co., Cambridge, 1862, p. 23.
2. Secchi A. Le Soleil (2nd Edition, Part II). Guathier-Villars, Paris, 1877.
3. Robitaille P.M. A thermodynamic history of the solar constitution – I: The journey to a gaseous Sun. *Progr. Phys.*, 2011, v. 3, 3–25.
4. Athay R.G. The Solar Chromosphere and Corona: Quiet Sun. D. Reidel Publishing Co., Boston, M.A., 1976.
5. Thomas R.N. and Athay R.G. Physics of the Solar Chromosphere. Interscience Publishers, New York, N.Y., 1961.
6. Bray R.J. and Loughhead R.E. The Solar Chromosphere. Chapman and Hall Ltd., London, U.K., 1974.
7. Ulmschneider P. The physics of the chromosphere and corona. In: *Lectures on Solar Physics* (H.M. Antia, A. Bhatnagar and R. Ulmschneider, Eds.), Springer, Berlin, 2003, p. 232–280.
8. Heinzel P. Understanding the solar chromosphere. In: *"Exploring the Solar System and the Universe"*, (V. Mioc, C. Dumitrache, N.A. Popescu, Eds.), American Institute of Physics, 2008, 238–244.
9. Woltjer L. A photometric investigation of the spicules and the structure of the chromosphere. *Bull. Astron. Inst. Netherlands*, 1954, v. 12, no. 454, 165–176.

10. Rush J.H. and Roberts W.O. Recent studies of chromospheric spicules. *Australian J. Phys.*, 1954, v. 7, 230–243.
11. Beckers J.M. Solar spicules. *Ann. Rev. Astron. Astrophys.*, 1972, v. 10, 73–100.
12. Lorrain P. and Koutchmy S. Two dynamical models for solar spicules. *Solar Phys.*, 1996, v. 165, 115–137.
13. Sterling A. Solar spicules: A review of recent models and targets for future observations. *Solar Phys.*, 2000, v. 196, 79–111.
14. Zaqarashvili T.V. and Erdélyi R. Oscillations and waves in solar spicules. *Space Sci. Rev.*, 2009, v. 149, 355–388.
15. Pasachoff J.M., Jacobson W.A. and Sterling A.C. Limb spicules from ground and from space. *Solar Phys.*, 2009, v. 260, 59–82.
16. Zirin H. The mystery of the chromosphere. *Solar Phys.*, 1996, v. 169, 313–326.
17. Robitaille P.M. A high temperature liquid plasma model of the Sun. *Progr. Phys.*, 2007, v. 1, 70–81 (also in arXiv: astro-ph/0410075).
18. Robitaille P.M. Liquid Metallic Hydrogen: A Building Block for the Liquid Sun. *Progr. Phys.*, 2011, v. 3, 60–74.
19. Robitaille P.M. Liquid Metallic Hydrogen II: A Critical Assessment of Current and Primordial Helium Levels in Sun. *Progr. Phys.*, 2013, v. 2, 35–47.
20. Robitaille J.C. and Robitaille P.M. Liquid Metallic Hydrogen III. Intercalation and Lattice Exclusion Versus Gravitational Settling and Their Consequences Relative to Internal Structure, Surface Activity, and Solar Winds in the Sun. *Progr. Phys.*, 2013, v. 2, 87–97.
21. Robitaille P.M. Commentary on the liquid metallic hydrogen model of the Sun. Insight relative to coronal holes, sunspots, and solar activity. *Progr. Phys.*, 2013, v. 2, L7–L9.
22. Kippenhahn R. and Weigert A. Stellar structure and evolution. Springer-Verlag, Berlin, 1990.
23. Bahcall J.N. and Pinsonneault M.H. Standard solar models, with and without helium diffusion, and the solar neutrino problem. *Rev. Mod. Phys.*, 1992, v. 64, no.4, 885–926.
24. Bachall J.N., Pinsonneault M.H. and Wasserburg G.J. Solar models with helium and heavy-element diffusion. *Rev. Mod. Phys.*, 1995, v. 67, no. 4, 781–808.
25. Robitaille P.M. The solar photosphere: Evidence for condensed matter. *Progr. Phys.*, 2006, v. 2, 17–21.
26. Tsiropoula G., Tziotsious K., Kontogiannis I., Madjarska M.S., Doyle J.G. and Suematsu Y. Solar fine-scale structures I. Spicules and other small-scale, jet-like events at the chromospheric level: Observations and physical parameters. *Space Sci. Rev.*, 2012, v. 169, 181–244.
27. De Pontieu B., McIntosh S.W., Hansteen V.H., Carlsson M., Schrijver C.J., Tarbell T.D., Title A.M., Shine R.A., Suematsu Y., Tsuneta S., Katsukawa Y., Ichimoto K., Shimizu T. and Nagata S. A tale of two spicules: The impact of spicules on the magnetic chromosphere. *Publ. Astron. Soc. Japan*, 2007, v. 59, 655–660.
28. Zhang Y.Z., Shirata K, Wang J.X., Mao X.J., Matsumoto T., Liu Y. and Su T.T. Revision of spicule classification. *Astrophys. J.*, 2012, v. 750, 16 (9 pages).
29. Menzel D.H. A Study of the Solar Chromosphere. *Publications of the Lick Observatory*, University of California Press, Berkeley, CA, v. 17, 1931.
30. Menzel D.H. and Cillié G.G. Hydrogen emission in the chromosphere. *Astrophys. J.*, 1937, v. 85, 88–106.
31. Russell H.N. On the composition of the Sun's atmosphere. *Astrophys. J.*, 1929, v. 70, 11–82.
32. Khanna S.N. and Jena P. Atomic clusters: Building blocks for a class of solids. *Phys. Rev. B*, 1995, v. 51, no. 19, 13705–13716.
33. Bačić Z. and Miller R.E. Molecular clusters: Structure and dynamics of weakly bound systems. *J. Phys. Chem.*, 1996, v. 100, 12945–12959.
34. Alivisatos A.P., Barbara P.F., Castleman A.W., Chang J., Dixon D.A., Klein M.L, McLendon G.L., Miller J.S., Ratner M.A., Rossky P.J., Stupp S.I., Thompson M.E. From molecules to materials: Current trends and future directions. *Adv. Materials*, 1998, v. 10, no. 16, 1297–1336.
35. Claridge S.A., Castleman A.W., Khanna S.N., Murray C.B., Sen A. and Weiss P.S. Cluster-assembled materials. *ACS Nano*, 2009, v. 3, no. 2, 244–255.
36. Castleman A.W. and Bowen K.H. Clusters: Structure, energetics, and dynamics of intermediate states of matter. *J. Chem. Phys.*, 1996, v. 100, no. 31, 12911–12944.
37. Castlemann A.W. and Khanna S.N. Clusters, superatoms, and building blocks of new materials. *J. Phys. Chem. C*, 2009, v. 113, 2664–2675.
38. Bergeron D.E., Castleman A.W., Morisato T. and Khanna S.N. Formation of $Al_{13}I^-$: Evidence for the superhalogen character of Al_{13}. *Science*, 2004, v. 304, no. 5667, 84–87.
39. Kronik L., Fromherz R., Ko E., Ganterför G. and Chelikowsky J.R. Highest electron affinity as a predictor of cluster anion structures. *Nature Materials*, 2002, v. 1, no. 1, 49–53.
40. Konig L., Rabin I, Schultze W. and Ertl G. Chemiluminescence in the agglomeration of metal clusters. *Science*, 1996, v. 274, no. 5291, 1353–1355.
41. Ievlev D., Rabin I., Schulze W. and Ertl G. Light emission in the agglomeration of silver clusters. *Chem. Phys. Let.*, 2000, v. 328, 142–146.
42. Phillips K.J.H., Feldman U. and Landi E. Ultraviolet and X-ray Spectroscopy of the Solar Atmosphere. Cambridge University Press, Cambridge, U.K., 2008.
43. Babcock H.W. The Paschen series of hydrogen line in the spectrum of the solar chromosphere. *Pub. Astron. Soc. Pacific*, 1932, v. 44, no. 261, 323–324.
44. Becker R. Molecular hydrogen emission from star-forming regions in the Large Magellanic Cloud. *The Messanger*, 1989, v. 56, 57–59.
45. Arimoto N., Sofue Y. and Tsujimoto T. CO-to-H_2 conversion factor in galaxies. *Publ. Astron. Soc. Japan*, 1996, v. 48, 275–284.
46. Bolatto A.D., Wolfire M. and Leroy A.K. The CO-to-H_2 conversion factor. arXiv:1301.3498v2 [astro-phGA] (March 8, 2013).
47. Spinrad H. Observation of stellar molecular hydrogen. *Astrophys. J.*, 1966, v. 145, 195–205.
48. Jaeggli S.A., Lin H. and Uitenbroek H. On molecular hydrogen formation and the magnetohydrostatic equilibrium of sunspots. *Astrophys. J.*, 2012, v. 745, 133 (16pages).
49. Jordan C., Bueckner G.E., Bartoe J.D.F., Sandlin G.D. and Vanhoosier M.E. Emission lines of H_2 in the extreme-ultraviolet solar spectrum. *Astrophys. J.*, 1978, v. 226, 687–697.
50. Bartoe J.D.F., Brueckner G.E., Nicolas K.R., Sandlin G.D., Vanhoosier M.E. and Jordan C. H_2 emission in the solar atmosphere. *Mon. Not. Roy. Astron. Soc.*, 1979, v. 187, 463–471.
51. Sandlin G.D., Bartoe J.D.F., Brueckner G.E., Tousey R. and Vanhoosier M.E. The high-resolution solar spectrum, 1175–1770Å. *Astrophys. J. Suppl. Ser.*, 1986, v. 61, 801–898.
52. Innes D.E. SUMER-Hinode observations of microflares: excitation of molecular hydrogen. *Astron. Astrophys.*, 2008, v. 481, no. 1, L41–L44.
53. Przybilla N. and Butler K. The solar hydrogen spectrum in non-local thermodynamic equilibrium. *Astrophys. J.*, 2004, v. 610, L61–L64.
54. Robitaille P.M. Stellar opacity: The Achilles' heel of the gaseous Sun. *Progr. Phys.*, 2011, v. 3, 93–99.
55. Unsöld A. Über die Struktur der Fraunhofersehen Linien und die quantitative Spektralanalyse der Sonnenatmosphäre. *Zeitschrift für Physik*, 1928, v. 46, no. 11–12, 765–781.
56. Wöhl H. On molecules in sunspots. *Solar Phys.*, 1971, v. 16, 362–372.

57. Sinha K. Molecules in the Sun. *Proc. Astron. Soc. Australia*, 1991, v. 9, 32–36.

58. Olmsted C.M. Sun-spot bands which appear in the spectrum of a calcium arc burning in the presence of hydrogen. *Astrophys. J.*, 1908, v. 27, 66–69.

59. Shaw B.L. Inorganic Hydrides. Pergamon Press, Oxford, 1967.

60. Athay R.G. Line broadening in chromospheric spicules. *Astrophys. J.*, 1961, v. 134, 756–765.

61. Zirker J.B. The solar H and K lines of ionized calcium. *Solar Phys.*, 1968, v. 3, 164–180.

62. Gulyaev R.A. and Livshits M.A. Width of $Ca^{+}H$ line in spicules. *Soviet Astron.*, 1966, v. 9, no. 4, 661–663.

63. Robitaille P.M. On solar granulations, limb darkening, and sunspots: Brief insights in remembrance of Father Angelo Secchi. *Progr. Phys.*, 2011, v. 3, 79–88.

LETTERS TO PROGRESS IN PHYSICS

The Liquid Metallic Hydrogen Model of the Sun and the Solar Atmosphere V. On the Nature of the Corona

Pierre-Marie Robitaille

Department of Radiology, The Ohio State University, 395 W. 12th Ave, Columbus, Ohio 43210, USA.
robitaille.1@osu.edu

The E-corona is the site of numerous emission lines associated with high ionization states (i.e. FeXIV-FeXXV). Modern gaseous models of the Sun require that these states are produced by atomic irradiation, requiring the sequential removal of electrons to infinity, without an associated electron acceptor. This can lead to computed temperatures in the corona which are unrealistic (i.e. ∼30–100 MK contrasted to solar core values of ∼16 MK). In order to understand the emission lines of the E-corona, it is vital to recognize that they are superimposed upon the K-corona, which produces a continuous spectrum, devoid of Fraunhofer lines, arising from this same region of the Sun. It has been advanced that the K-corona harbors self-luminous condensed matter (Robitaille P.M. The Liquid Metallic Hydrogen Model of the Sun and the Solar Atmosphere II. Continuous Emission and Condensed Matter Within the Corona. *Progr. Phys.*, 2013, v. 3, L8–L10; Robitaille P.M. The Liquid Metallic Hydrogen Model of the Sun and the Solar Atmosphere III. Importance of Continuous Emission Spectra from Flares, Coronal Mass Ejections, Prominences, and Other Coronal Structures. *Progr. Phys.*, 2013, v. 3, L11–L14). Condensed matter can possess elevated electron affinities which may strip nearby atoms of their electrons. Such a scenario accounts for the high ionization states observed in the corona: condensed matter acts to harness electrons, ensuring the electrical neutrality of the Sun, despite the flow of electrons and ions in the solar winds. Elevated ionization states reflect the presence of materials with high electron affinities in the corona, which is likely to be a form of metallic hydrogen, and does not translate into elevated temperatures in this region of the solar atmosphere. As a result, the many mechanisms advanced to account for coronal heating in the gaseous models of the Sun are superfluous, given that electron affinity, not temperature, governs the resulting spectra. In this regard, the presence of highly ionized species in the corona constitutes the thirty-first line of evidence that the Sun is composed of condensed matter.

> *In order to explain the occurrence of the dark lines in the solar spectrum, we must assume that the solar atmosphere incloses a luminous nucleus, producing a continuous spectrum, the brightness of which exceeds a certain limit. The most probable supposition which can be made respecting the Sun's constitution is, that it consists of a solid or liquid nucleus, heated to a temperature of the brightest whiteness, surrounded by an atmosphere of somewhat lower temperature.*
>
> Gustav Robert Kirchhoff, 1862 [1]

Superimposed on the continuous spectrum of the inner K-corona are emission lines, including one at 5303.3 Å, the famous line from coronium, first discovered by Harkness and Young [2, 3], photographed by Evershed [4], and eventually identified as FeXIV by Bengt Edlén [5–7]. Walter Grotian suggested that this line originated from highly ionized atoms, supported by early reports of similar findings from Bengt Edlén in such atoms [5–8]. The wonderful story of coronium [5, 6], along with the roles played by Walter Grotian and Bengt Edlén has been presented by Edward A. Milne [7].

Milne's account provides a key fact relative to coronium: the formation of FeXIV requires energy in the soft X-ray range of the electromagnetic spectrum [7], but the Sun emits very few of these rays. As such, how does one produce ions with such elevated ionization states in the corona?

Today, X-ray spectroscopy reveals that the Sun can produce emission lines from ions with ionization states as high as FeXXV [9]. Within the context of the gaseous models [10–12], the formation of such species calls for the removal of electrons from electronic shells to infinity, requiring energies associated with temperatures of ∼30 MK [9, p. 26]. It has also been postulated that superhot thermal components ($>10^8$ K) can be generated above the limb in association with some flares [13] and radio studies initially called for temperatures of 10^8–10^{10} K in the corona [14, p. 128].

In 2000, the Bastille Day flare produced FeXII lines, but with a spine emmiting FeXXIV lines [9, p. 19]. If such findings are to be explained within the context of a gaseous solar model [10–12], it is not surprising that extreme temperatures must be invoked. A gaseous Sun has no other means of pro-

ducing highly ionized species.

At the same time, the extreme temperatures currently associated with the corona must be viewed with caution, given that the core of the Sun has been postulated to harbor temperatures of only ~16 MK [10, p. 9]. In addition, it is claimed that the energy source driving such extremes in temperature *"must be magnetic since all the other possible sources are completely inadequate"* [13]. Such statements, and the computed temperatures from which they stem, directly reflect the shortcomings of the gaseous solar models [10–12]. The need to explain the synthesis of highly ionized ions in the corona within a gaseous context is so acute that numerous schemes have been advanced to heat the chromosphere and corona [15, 16]. Ulmschneider states that *"The chromosphere and corona are thus characterized as layers which require large amounts of mechanical heating"* [15, p, 235] and further *"To clarify the zoo of coronal heating processes much further work remains to be done"* [15, p. 278].

Since the corona must be excessively hot to produce such ions in a gaseous context, the continuous spectrum of the K-corona has been dismissed as a strange artifact, produced by electronic scattering of photospheric light [17]. Otherwise, the coronal continuous spectrum would be indicating that *apparent coronal temperatures are no warmer than those of the photosphere*. It would be impossible for the gaseous models [10–12] to account for the presence of highly ionized species within the outer solar atmosphere. Consequently, sufficient electron densities are inferred to exist in the corona to support the idea that the spectrum of the K-corona is being produced by the scattering of photospheric light: *"The reason we see the corona in white, or integrated, light is that the photospheric light is scattered by coronal electrons: we see the light that does not get through but is scattered towards us. This scattered light is about 10^{-6} as intense as the photospheric light, which means it has been scattered by 10^{19} electrons; these are distributed along a path about equal to the diameter of the sun, or 1.4 x 10^{11} cm, so the average coronal density close to the surface must be 10^{8} electrons/cm^3"* [14, p. 75]. Much like the solar surface [18], the relevance of a thermal spectrum in the K-corona has been rejected as little more than an optical illusion [17].

In the end, all extreme temperatures obtained from line emission should be dismissed as erroneous. Discovery of FeXXV within X-ray flares suggests that we do not properly understand the formation of emission lines with high ionization levels in the corona. Current temperature estimates are flirting with violations of both the first and second laws of thermodynamics: it is difficult to conceive that localized temperatures within flares and the corona could greatly *exceed* the temperature of the solar core.

Instead, line emission spectra from highly ionized ions might best be viewed as direct evidence that materials with elevated electron affinities exist within the corona. Such a solution can be readily associated with the condensed nature of the Sun [19–23].

In this regard, the continuous spectrum of the K-corona must be regarded as genuine [17]. The slight reddening of the K-corona, reported long ago by Allen [24], indicates that *apparent* coronal temperatures are gently decreasing with increasing distance above the solar surface. The corona seems to contain condensed matter of the same nature as found on the photosphere, since the spectrum of the K-corona, though devoid of Fraunhofer lines, is essentially identical to that produced by the solar surface [18]. This proposal is compelling, given that the Sun is expelling material into its corona which is also known to emit continuous visible spectra [25].

By extension, *apparent* coronal temperatures, which are likely to represent vibrational lattice phenomena [26–29], can be no greater than those found on the surface of the Sun. Therefore, contrary to popular scientific belief [15, 16], the corona of the Sun is *not* being heated. Rather, free atoms in the corona are being stripped of their electrons, as they interact with condensed matter which possesses much higher electron affinity. Neutral atoms have limited electron affinities, but molecules can have higher values.* However, condensed matter can develop enormous attractive forces for electrons.

This lesson is well manifested on Earth, as lightning attempts to equalize charge imbalance between separate regions of condensed matter [31–33]. Typically, lightning forms in clouds containing solid or liquid water particles. But it can also occur *"above volcanoes, in sandstorms, and in nuclear explosions"* [33, p. 67]. Usually, lightning forms between different cloud regions, or between clouds and the Earth's surface [31–33]. Lightning represents the longest standing example of the power of electron affinity in condensed matter. In this respect, while temperatures in the tens of thousands of degrees could be inferred from Hα line analysis during lightning activity,† scientists do not claim that the atmosphere of the Earth exists at these temperatures.

Thunderhead clouds can generate substantial steady electric fields on the order of 100 kVm^{-1} [33, p. 494]. Such fields have have been associated with runaway electrons capable of generating X-rays with energies of 100 KeV or more [33, p. 493-495]. Nonetheless, these energies are not translated into associated temperatures, as values in excess of 10^9 K would be derived. Still, for the purpose of this discussion, it is important to note that the presence of condensed matter in the atmosphere of the Earth can lead to amazing phenomena, when electric potentials are eliminated through charge transfer.

The author has advanced that the corona of the Sun is filled with sparse remnants of liquid metallic hydrogen [18] which have been expelled from the body of the Sun [25]. Such material is expected to have a highly conductive nature

*Of the elements, chlorine has the highest electron affinity at ~3.6 eV, calcium has the lowest value at ~0.02 eV; molecular RuF_6 has a value of ~7.5 eV [30]

†Peak temperatures of ~35,000 K have been reported [33, p. 163]

and could be used to harvest electrons from the corona, helping to ensure the continued neutrality of the solar body and solar winds. The presence of metallic hydrogen in the corona may then promote, through its elevated electron affinity, the creation of highly ionized species.

For instance, when iron comes in contact with metallic hydrogen, MH, it could initially form an activated complex, MH–Fe*, MH + Fe $\rightarrow$ MH–Fe*. This excited complex then relaxes by capturing n electrons from the iron atom. This could be accomplished with the simultaneous ejection of an activated iron species, Fe^{+n*}, leading to the following reaction: MH–Fe$^* \rightarrow$ MH–$n\bar{e}$ + Fe^{+n*}. The resulting excited iron could then relax back to the ground state through line emission, $Fe^{+n*} \rightarrow Fe^{+n}$ + hv. Depending on the local electron affinity of metallic hydrogen, n could range from single digits to ~25 [9] in the case of iron. A similar process could be invoked to create the other highly ionized species of the corona. In this regard, it is interesting to note that most of the ions observed in the solar *"XUV spectrum are principally those with one or two valence electrons"* remaining [14, 173].

In this scenario, the electron affinity of metallic hydrogen in the outer atmosphere responds to charge imbalances, either in the corona itself or on the surface of the Sun, by capturing electrons locally. Metallic hydrogen in the corona thereby acts as a conductive medium surrounding the solar body, constantly ensuring overall charge neutrality for the Sun. The arrangement of coronal steamers is highly suggestive of such a role from these objects, though all coronal structures might be involved in the recapture of electrons from the outer solar atmosphere.

Outstanding images of the corona have been obtained using spectroscopic lines from highly ionized iron (e.g. FeX–FeXIV) [34–37]. The presence of FeX–FeXIV throughout the solar atmosphere strengthens the concept that interactions between atoms and metallic hydrogen in the corona act to maintain neutrality on the Sun by producing highly ionized atoms throughout this region.

Moreover, flare studies indicate that coronal structures can display highly organized local electron affinities. As mentioned earlier, the TRACE team has produced a flare image where central spine structures produce line emission from FeXXIV and CaXVII, while the exterior of the flare emits in FeXII [9, p. 19]. Such images would be nearly impossible to explain in the context of a gaseous model of the Sun. Instead, organized structures within the corona and its components are strongly supportive of the idea that the Sun is comprised of condensed matter.

In closing, the liquid metallic hydrogen model of the Sun [19–23] provides an elegant solution for the production of highly ionized species in the corona. The wide variety of oxidation states can be simply obtained by invoking regions of varying electron affinity within the condensed structures that comprise the corona. The complete, or significant, removal of electrons from atoms can be explained using a single interaction, namely the temporary contact between atoms and metallic hydrogen.

The production of such ions in the gaseous models [10–12] requires the repeated ejection of electrons from their orbitals in a multistage process, whereby up to two dozen events must logically follow one another. Studies indicate the existence of species such as C^{+6}, Fe^{+14} and Fe^{+16} in the solar wind [38, p. 114]. Such ions require multiple steps for production in a gaseous context [10–12] and would be the result of random processes.

Conversely, the synthesis of highly ionized atoms requires but a single step in the liquid metallic hydrogen model [19–23]. The generation of such ions is no longer a random act, but rather a direct manifestation of the function of the corona, *facilitation of electron capture in the outer atmosphere of the Sun in order to preserve solar neutrality*. The production of highly ionized species throughout the corona therefore constitutes the thirty-first line of evidence that the Sun is composed of condensed matter.

Dedication

Dedicated to the poor, who sleep, nearly forgotten, under the light of the Southern Cross.

Submitted on: May 1, 2013 /Accepted on: May 2, 2013
First published online on: May 13, 2013

References

1. Kirchhoff G. The physical constitution of the Sun. In: *Researches on the Solar Spectrum and the Spectra of the Chemical Elements*. Translated by H.E. Roscoe, Macmillan and Co., Cambridge, 1862, p. 23.
2. Hufbauer K. Exploring the Sun: Solar Science since Galileo. The Johns Hopkins University Press, Baltimore, 1991, p. 112–114.
3. Dick S. Sky and Ocean Joined: The U.S. Naval Observatory 1830–2000. Cambridge University Press, Cambridge, 2003, p. 196-205.
4. Evershed J. Wave-length determinations and general results obtained from a detailed examination of spectra photographed at the solar eclipse of January 22, 1898. *Phil. Trans. Roy. Soc. London*, 1901, v. 197, 381–413.
5. Claridge G.C. Coronium. *J. Roy. Astron. Soc. Canada*, 1937, v. 31, no. 8, 337–346.
6. Unsigned. Origin of the coronium lines. *Nature*, 1942, v. 150, no. 3817, 756–759.
7. Milne A.E. Presidential Address – Award of the Gold Medal to Professor Bengt Edlén. *Mon. Not. Roy. Astron. Soc.*, 1945, v. 105, 138–145.
8. Grotian W. über die intensitätsverteilung des kontinuierlichen spektrums der inneren korona. *Zeitschrift fur Astrophysik*, 1931, v. 3, 199–226.
9. Phillips K.J.H., Feldman U. and Landi E. Ultraviolet and X-ray Spectroscopy of the Solar Atmosphere. Cambridge University Press, Cambridge, 2008.
10. Kippenhahn R. and Weigert A. Stellar structure and evolution. Springer-Verlag, Berlin, 1990.
11. Bachall J.N., Pinsonneault M.H. and Wasserburg G.J. Solar models with helium and heavy-element diffusion. *Rev. Mod. Phys.*, 1995, v. 67, no. 4, 781–808.
12. Robitaille P.M. The solar photosphere: Evidence for condensed matter. *Progr. Phys.*, 2006, v. 2, 17–21.

13. Priest E.R. Solar flare theory and the status of flare understanding. *High Energy Solar Physics: Anticipating HESSI (R. Ramaty and N. Mandzhavidze, Eds.), ASP Conf. Ser.*, 2000, v. 206, 13–26.

14. Zirin H. The Solar Atmosphere. Blaisdell Publishing Company, Waltham, M.A., 1966.

15. Ulmschneider P. The physics of the chromosphere and corona. In: *Lectures on Solar Physics* (H.M. Antia, A. Bhatnagar and R. Ulmschneider, Eds.), Springer, Berlin, 2003, p. 232–280.

16. Dwivedi B.N. The solar corona. In: *Lectures on Solar Physics* (H.M. Antia, A. Bhatnagar and R. Ulmschneider, Eds.), Springer, Berlin, 2003, p. 281–298.

17. Robitaille P.M. The Liquid Metallic Hydrogen Model of the Sun and the Solar Atmosphere II. Continuous Emission and Condensed Matter Within the Corona. *Progr. Phys.*, 2013, v. 3, L8–L10.

18. Robitaille P.M. On the Presence of a Distinct Solar Surface: A Reply to Hervé Faye. *Progr. Phys.*, 2011, v. 3, 75–78.

19. Robitaille P.M. A high temperature liquid plasma model of the Sun. *Progr. Phys.*, 2007, v. 1, 70–81 (also in arXiv: astro-ph/0410075).

20. Robitaille P.M. Liquid Metallic Hydrogen: A Building Block for the Liquid Sun. *Progr. Phys*., 2011, v. 3, 60–74.

21. Robitaille P.M. Liquid Metallic Hydrogen II: A Critical Assessment of Current and Primordial Helium Levels in Sun. *Progr. Phys*., 2013, v. 2, 35–47.

22. Robitaille J.C. and Robitaille P.M. Liquid Metallic Hydrogen III. Intercalation and Lattice Exclusion Versus Gravitational Settling and Their Consequences Relative to Internal Structure, Surface Activity, and Solar Winds in the Sun. *Progr. Phys*., 2013, v. 2, 87-97.

23. Robitaille P.M. Commentary on the liquid metallic hydrogen model of the Sun. Insight relative to coronal holes, sunspots, and solar activity. *Progr. Phys*., 2013, v. 2, L7–L9.

24. Allen C.W. The spectrum of the corona at the eclipse of 1940 October 1. *Mon. Not. Roy. Astron. Soc.*, 1946, v. 106, 137–150.

25. Robitaille P.M. The Liquid Metallic Hydrogen Model of the Sun and the Solar Atmosphere III. Importance of Continuous Emission Spectra from Flares, Coronal Mass Ejections, Prominences, and Other Coronal Structures. *Progr. Phys*., 2013, v. 3, L11–L14.

26. Robitaille P.M. On the validity of Kirchhoff's law of thermal emission. *IEEE Trans. Plasma Sci.*, 2003, v. 31, no. 6, 1263–1267.

27. Robitaille P.M. A critical analysis of universality and Kirchhoff's law: A return to Stewart's law of thermal emission. *Prog. Phys.*, 2008, v. 3, 30–35.; arXiv:0805.1625.

28. Robitaille P.M. Blackbody radiation and the carbon particle. *Progr. Phys.*, 2008, v. 3, 36–55.

29. Robitaille P.M. Kirchhoff's Law of Thermal Emission: 150 years. *Progr. Phys.*, 2009, v. 4, 3–13.

30. Hayes W.M (Editor-in-Chief), CRC Handbook of Chemistry and Physics, 93rd Edition, Internet version 2013, 10:147–10:162.

31. Uman M.A. Lightning. Dover Publicatons. New York, N.Y., 1984.

32. Uman M.A. The Lightning Discharge (International Geophysics Series – Vol. 39), Academic Press, Inc., New York, N.Y., 1987.

33. Rakov V.A. and Uman M.A. Lightning: Physics and Effects. Cambridge University Press, Cambridge, U.K., 2003.

34. Wood B.E., Karovska M., Cook J.W., Brueckner G.E., Howard R.A., Korendyke C.M. and Socker D.G. Search for brightness variations in FeXIV coronagraph observations of the quiescent solar corona. *Astrophys. J.*, 1998, v. 505, 432–442.

35. Habbal S.R., Druckmüller M., Morgan H., Daw A., Johnson J., Ding A., Arndt M., Esser R., Rušin V. and Scholl I. Mapping the distribution of electron temperature and Fe charge states in the corona with total solar eclipse observations. *Astrophys. J.*, 2010, v. 708, 1650–1662.

36. Habbal S.R., Druckmüller M., Morgan H., Scholl I., Rušin V., Daw A., Johnson J. and Arndt M. Total solar eclipse observations of hot prominence shrouds. *Astrophys. J.*, 2010, v. 719, 1362–1369.

37. Habbal S.R., Morgan H. and Druckmüller M. A new view of coronal structures: Implications for the source and acceleration of the solar wind – First Asia-Pacific Solar Physics Meeting. *ASI Conf. Ser.*, 2011, v. 2, 259–269.

38. Gosling J.T. The solar wind in *Encyclopedia of the Solar System, 2nd Edition*, (L.A. McFadden, P.R. Weissman and T.V. Johnson, Eds.), Academic Press, San Diego, C.A., 2007, 99–116.

LETTERS TO PROGRESS IN PHYSICS

The Liquid Metallic Hydrogen Model of the Sun and the Solar Atmosphere VI. Helium in the Chromosphere

Pierre-Marie Robitaille

Department of Radiology, The Ohio State University, 395 W. 12th Ave, Columbus, Ohio 43210, USA.
robitaille.1@osu.edu

Molecular hydrogen and hydrides have recently been advanced as vital agents in the generation of emission spectra in the chromosphere. This is a result of the role they play in the formation of condensed hydrogen structures (CHS) within the chromosphere (P.M. Robitaille. The Liquid Metallic Hydrogen Model of the Sun and the Solar Atmosphere IV. On the Nature of the Chromosphere. *Progr. Phys.*, 2013, v. 3, 15–21). Next to hydrogen, helium is perhaps the most intriguing component in this region of the Sun. Much like other elements, which combine with hydrogen to produce hydrides, helium can form the well-known helium hydride molecular ion, HeH^+, and the excited neutral helium hydride molecule, HeH^*. While HeH^+ is hypothesized to be a key cosmological molecule, its possible presence in the Sun, and that of its excited neutral counterpart, has not been considered. Still, these hydrides are likely to play a role in the synthesis of CHS, as the He I and He II emission lines strongly suggest. In this regard, the study of helium emission spectra can provide insight into the condensed nature of the Sun, especially when considering the 10830 Å line associated with the $2^3P \rightarrow 2^3S$ triplet state transition. This line is strong in solar prominences and can be seen clearly on the disk. The excessive population of helium triplet states cannot be adequately explained using the gaseous models, since these states should be depopulated by collisional processes. Conversely, when He-based molecules are used to build CHS in a liquid metallic hydrogen model, an ever increasing population of the 2^3S and 2^3P states might be expected. The overpopulation of these triplet states leads to the conclusion that these emission lines are unlikely to be produced through random collisional or photon excitation, as required by the gaseous models. This provides a significant hurdle for these models. Thus, the strong $2^3P \rightarrow 2^3S$ lines and the overpopulation of the helium triplet states provides the thirty-second line of evidence that the Sun is comprised of condensed matter.

> *In order to explain the occurrence of the dark lines in the solar spectrum, we must assume that the solar atmosphere incloses a luminous nucleus, producing a continuous spectrum, the brightness of which exceeds a certain limit. The most probable supposition which can be made respecting the Sun's constitution is, that it consists of a solid or liquid nucleus, heated to a temperature of the brightest whiteness, surrounded by an atmosphere of somewhat lower temperature.*
>
> Gustav Robert Kirchhoff, 1862 [1]

Estimates of solar helium abundances have varied widely over the years. For instance, *"different methods and different data sets give values ranging from 20% to 40% of the Sun's mass"* [2, p. 381]. 'Primordial' helium levels strongly guide all solar helium abundance determinations, as the amount of helium in the stars is said to be closely correlated with the synthesis of this element soon after the Big Bang [3–6]. Helium abundances currently act as one of the "Great Pillars" of cosmology (see [7] for detailed discussion). As such, any attempt to alter accepted helium levels within the Sun has great implications throughout astrophysics.

Recently, the author has reviewed the determination of solar helium abundances and reached the conclusion that these levels are likely to have been overstated [7]. The most prudent outlook remains that the Sun, like the visible universe, is composed primarily of hydrogen, as first outlined by Cecilia Payne [8]. In this regard, Robitaille and Robitaille have highlighted that the solar body is apt to be excluding helium from its interior [9]. It is well-known that this element can be expelled from the Sun during periods of elevated solar activity with widely varying quantities observed in the associated solar wind [10–14]. As a result, it is unlikely that the Sun is harboring much helium [7, 9]. Significant levels of helium above the photosphere merely represent eons of helium synthesis in a hydrogen based Sun. It can be hypothesized that since helium cannot re-enter the Sun once expelled, it slowly accumulates as a gas within the chromospheric region.

In his classic textbook, *"Astrophysics of the Sun"*, Harold Zirin emphasized that the helium D3 line can be enhanced

more than 20 fold, as viewing moves from the center of the solar disk to just beyond the limb, displaying *"a sharp spike"* [15, p. 199–200]. He outlined that this emission *"comes from a low thin layer"* [15, p. 198]. Similarly, Zirin states that the triplet He I transition at 10830 Å is barely visible on the disk, but almost as strong as Hα at the limb [15, p. 199–200]. Moreover, he adds that the λ1640 line is known to increase in intensity at least fifteen times near the limb, while lines from neutral helium are enhanced 50 fold [15, p. 199–200]. Since helium emission peaks at ~1200 km above the photosphere, these findings strongly suggest that the element is floating in a cloud lying several hundred kilometers above the surface, although He remains sparse over coronal holes [15, p. 198].

At the same time, though relatively faint, helium lines are present in the spicules [16]. Since chromospheric structures, like spicules, have been hypothesized to be the site of hydrogen condensation in the solar atmosphere, it is important to understand why helium emission lines are associated with such objects.

Based on the chemiluminescence observed when silver clusters condense [17], the author has recently stated that all emission lines originating in the flash spectrum are a direct consequence of condensation in this region of the Sun [18]. By necessity, these exothermic condensation reactions involve the ejection of an excited atomic species from the condensate which can then relax back to a lower energy level through the emission of a photon. For instance, the Ca II emission, which is so typical of the chromospheric spectrum, has been hypothesized to involve the reaction of CaH and a condensed hydrogen structure, CHS [18], to create an excited complex, CHS + CaH $\rightarrow$ CHS–HCa*. This step is then followed by the exothermic expulsion of an excited Ca II ion, CHS–HCa $\rightarrow$ CHS–H + Ca^{+*}, and later by line emission from Ca II*, Ca^{+*} $\rightarrow$ Ca^{+} + $h\nu$. Similar reactions have been invoked for all the hydrides present on the Sun [18]. The most significant of these take place using molecular hydrogen, and this explains the prevalence of strong emission lines from this element in the chromosphere. In order to account for the He I and He II emission lines associated with the flash spectrum, a directly analogous scenario must be invoked, which this time requires a helium hydride molecular species.

Many charged molecular ions of helium have been studied. The most famous, helium hydride, HeH^{+}, *is ubiquitous in discharges containing hydrogen and helium.*" [19]. This molecular cation was first discovered experimentally in 1925 by Hogness and Lunn [20]. It has been the focus of extensive spectroscopic studies [21, 22] and also postulated to play a key role in chemical astrophysics [23–25]. Wolfgang Ketterle (Nobel Prize, Physics, 2001) was the first to obtain its spectroscopic lines [26, 27]. The molecule has a bond distance of 0.77Å and a dissociation energy of ~44.6 kcal mol^{-1} [19]. Although it exists only in the gas phase, its Brønsted acidity should be extremely powerful. As a result, the hydrogen hydride cation should have a strong tendency to donate a proton,

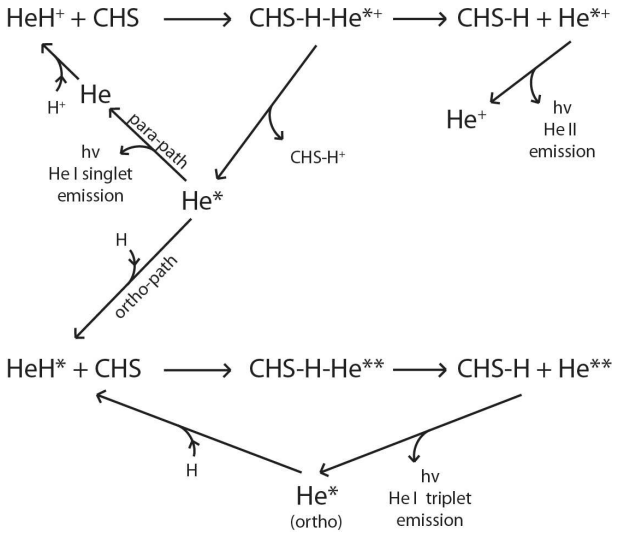

Fig. 1: Schematic representation of possible pathways involved when the helium hydride ion, HeH^{+}, or the excited helium hydride molecule, HeH*, react with condensed hydrogen structures, CHS, in the chromosphere of the Sun. The pathways presented can account for all emission lines observed from He I and He II. Note in this scheme that excited helium, He*, is being produced initially through the interaction of HeH^{+} with CHS. This excited helium, He*, if it assumes the triplet state (orthohelium — electrons in the same orientation: spin up/up or down/down), will then become trapped in the excited state. This triplet helium can then be used repeatedly, in cyclic fashion, to condense hydrogen atoms onto chromospheric structures, CHS (as shown in the lower half of the figure). Alternatively, if excited helium He* is initially produced in the singlet state (parahelium — electrons in different orientation: spin up/down), emission can immediately occur generating the singlet lines from He I. This scheme accounts for the strong triplet He I transition at 10830 Å observed in the flash spectrum of the chromosphere. Unlike the situation in the gas models, random collisional or photon excitations are not invoked to excite the helium atoms. As a result, de-excitation processes would also be absent, helping to ensure the buildup of triplet state orthohelium in this model.

without the concerted transfer of an electron.

In the chromosphere, the interaction between the helium hydride ion, HeH^{+}, and condensed hydrogen structures, CHS [18], could lead to an array of reactions as outlined in Fig. 1.

The simple combination of HeH^{+} and CHS could form an activated complex: CHS + HeH^{+} $\rightarrow$ CHS–H–He^{+*}. Exothermic expulsion of an excited helium ion, He^{+*}, could follow with full transfer of a proton and an electron to the condensed hydrogen structure: CHS–HHe^{+*} $\rightarrow$ CHS–H + He^{+*}. The resulting He^{+*} would be able to relax back to a lower energy state through emission, He^{+*} $\rightarrow$ He^{+} + $h\nu$, leading to the well known He II lines in the chromosphere (see Fig. 1).

Alternatively, when HeH^{+} reacts with CHS, it could lead initially to the same condensation adduct, CHS–H–He^{+*}, but this time, exothermic expulsion of an excited helium atom

could follow (see Fig. 1). Since HeH^+ should be a strong Brønsted acid, the transfer of a proton to the CHS could occur without electron transfer: CHS–HHe^{+*} → CHS–H^+ + He^*. This leads to several phenomena.

First, the relaxation of an excited helium atom, does not involve the same processes which occur in the helium ion. This is because the He^{+*} possesses only a single electron. As such, the electron in He^{+*} can simply relax back down to any lower energy level, including the ground state, giving the well-known He II lines on the Sun.

Conversely, because an excited helium atom contains two electrons, the possible fate of this species is more complicated. Since one of the electrons has not been excited, it remains in the lowest energy state, with a given spin, either up or down. The excited electron is only allowed by selection rules to return to the ground state, if and only if, its spin is opposed to that of the ground state electron.

If the ground state electron is 'spin down', then the excited electron can make the transition back to the ground state if it is 'spin up'. Helium in this case is known as *parahelium* (or singlet helium), emphasizing that its two electrons have spins with opposite orientation. The singlet He^* would simply relax back to the ground state, given rise to the emission lines from the neutral atom, He I, $He^* \rightarrow He + h\nu$. Likewise, if the ground state electron is 'spin up', the excited electron must be 'spin down' to enable the transition, $He^* \rightarrow He + h\nu$, again producing the identical He I lines from singlet state parahelium.

However, if the two electrons of He^* have the same spin (both up or both down), then the excited electron cannot relax back to the ground state. It remains *trapped* in the excited state. Helium in this case is known as *orthohelium* (or triplet state helium), emphasizing that its two electrons have spins with the same orientation. It is the reactions of orthohelium which are of particular interest in this work, as their existence is elegantly accounted for through the condensation of hydrogen [18], as described below.

Since orthohelium is trapped in the excited triplet state, it has an opportunity to once again react with hydrogen, as displayed in the lower portion of Fig. 1. Wolfgang Ketterle has demonstrated that excited helium hydride also exists [28, 29]. Therefore, given a lack of relaxation, triplet He^* could capture a hydrogen atom, forming neutral excited helium hydride: $He^* + H \rightarrow HeH^*$. This species could once again react with CHS [18], but this time forming a doubly activated complex: CHS + $HeH^* \rightarrow$ CHS–H–He^{**}. The net transfer of a hydrogen atom in this case leads to release from the CHS of doubly excited helium.* When this occurs, the He^{**} atom is now able to relax, as the excited electron which is now in the 2p or 3s orbital, undergoes a transition down to the 2s orbital. The $2^3P \rightarrow 2^3S$ transition is associated with the strong triplet He I line at 10830 Å observed in the prominences and on the disk of the Sun [30, p. 95]. Alternatively, a $3^3P \rightarrow 2^3S$ transition produces the triplet He I line at 3890 Å [30, p. 95].

As illustrated in Fig. 1, once the doubly excited helium atom has partially relaxed to regenerate orthohelium, it can react once again with atomic hydrogen, leading to the renewed synthesis of excited helium hydride, HeH^*. A cyclic pathway has been created, wherein triplet hydrogen is preserved and continuously working to assist in the resynthesis of condensed hydrogen structures, as the Sun recaptures any atomic hydrogen lost to its atmosphere.

Importantly, the entire process is being 'primed' through the use of a single HeH^+ molecular ion and the intial transfer of a single proton to the CHS. This feature is noteworthy, since true condensation requires the transfer of electrons and protons to the chromospheric structures. In this regard, the generation of Ca II emission lines from analogous condensations of calcium hydride, involves the transfer of two electrons per hydrogen atom [18]. Such parallel reactions could help to ensure that overall charge balance in the building of condensed hydrogen structures can be maintained.

In the end, this approach holds many advantages over the random processes invoked by the gaseous models of the Sun in order to account for line emission in the chromosphere. All line emission in the chromosphere become directly associated with ordered reactions, whose product, CHS, are vital to preserving the solar mass. The Sun does not simply eject hydrogen into its atmosphere, without any hope of regaining these atoms. Rather, in the chromosphere, hydrogen atoms are constantly being recaptured through hydride based reactions. The triplet state of orthohelium, so strongly manifested within prominences and in the chromospheric emission spectrum, becomes not an incidental artifact, but rather, a necessary and direct manifestation that organized chemical reactions are taking place within the chromosphere. As such, the existence of this abundant orthohelium and the strong emission lines which it produces can be said to constitute the thirty-second line of evidence that the Sun is comprised of condensed matter.

Acknowledgment

Luc Robitaille is acknowledged for figure preparation.

Dedication

Dedicated to the poor, who sleep, nearly forgotten, under the light of the Southern Cross.

Submitted on: May 14, 2013 / Accepted on: May 16, 2013
First published online on: May 31, 2013

References

1. Kirchhoff G. The physical constitution of the Sun. In: *Researches on the Solar Spectrum and the Spectra of the Chemical Elements*. Translated by H.E. Roscoe, Macmillan and Co., Cambridge, 1862, p. 23.

*We can assume that the ground state electron remains stationary, but that the initially excited electron has now been transferred to an even higher atomic orbital. Alternatively, both electrons could be excited, but this case will not be considered.

2. Bhatnagar A. and Livingston W. Fundamentals of Solar Astronomy (World Scientific Series in Astronomy and Astrophysics – Vol. 6), World Scientific, New Jersey, 2005.
3. Peebles P.J.E. Primordial helium abundance and the primordial fireball. II. *Astrophys. J.*, 1966, v. 146, 542–552.
4. Danzinger I.J. The cosmic abundance of helium. *Ann. Rev. Astron. Astrophys.*, 1970, v. 8, 161–178.
5. Izotov Y.I. and Thuan T.X. The primordial abundance of 4He revisited. *Astrophys. J.*, 1998, v. 500, 188–216.
6. Olive K.A., Steigman G. and Walter T.P. Primordial nucleosynthesis: Theory and observations. *Phys. Rep.*, 2000, v. 333-334, 389–407.
7. Robitaille P.M. Liquid Metallic Hydrogen II: A critical assessment of current and primordial helium levels in Sun. *Progr. Phys.*, 2013, v. 2, 35–47.
8. Payne C.H. The relative abundances of the elements. Stellar Atmospheres. Harvard Observatory Monograph no. 1 (Harlow Shapley, Editor), Harvard University Press, Cambridge, MA, 1925 (reprinted in part in Lang K.R. and Gingerich O. A source book in astronomy and astrophysics, 1900–1975, Harvard University Press, Cambridge, MA, 1979, p. 245–248).
9. Robitaille J.C. and Robitaille P.M. Liquid metallic hydrogen III. Intercalation and lattice exclusion versus gravitational settling and their consequences relative to internal structure, surface activity, and solar winds in the Sun. *Progr. Phys.*, 2013, v. 2, 87–97.
10. Robbins D.E., Hundhausen A.J. and Bame S.J. Helium in the solar wind. *J. Geophys. Res.*, 1970, v. 75, no. 7, 1178-1187.
11. Bame S.J., Asbridge J.R., Feldman W.C. and Gosling J.T. Evidence for a structure-free state at high solar wind speeds. *J. Geophys. Res.*, 1977, v. 82, no. 10, 1487–1492.
12. Borrini G., Gosling J.T., Bame S.J. and Feldman W.C. Helium abundance enhancements in the solar wind. *J. Geophys. Res*, 1982, v. 87, no. A9, 7370–7378.
13. Aellig M.R., Lazarus A.J. and Steinberg J.T. The solar wind helium abundance: Variations with wind speed and solar cycle. *Geophys. Res. Let.*, 2001, v. 28, no. 14, 2767–2770.
14. Kasper J.C., Stevens M.L., Lazarus A.J. and Ogilvie K.W. Solar wind and helium abundance as a function of speed and heliographic latitude: Variation through a solar cycle. *Astrophys. J.*, 2007, v. 660, 901–910.
15. Zirin H. Astrophysics of the Sun. Cambridge University Press, Cambridge, U.K., 1988.
16. Zirin H. The mystery of the chromosphere. *Solar Phys.*, 1996, v. 169, 313–326.
17. Konig L., Rabin I, Schultze W. and Ertl G. Chemiluminescence in the agglomeration of metal clusters. *Science*, 1996, v. 274, no. 5291, 1353–1355.
18. Robitaille P.M. The Liquid Metallic Hydrogen Model of the Sun and the Solar Atmosphere IV. On the Nature of the Chromosphere. *Progr. Phys.*, 2013, v. 3, L15–L21.
19. Grandinetti F. Helium chemistry: A survey of the role of the ionic species. *Inter. J. Mass Spectrom.*, 2004, v. 237, 243–267.
20. Hogness T.R. and Lunn E.G. The ionization of hydrogen by electron impact as interpreted by positive ray analysis. *Phys. Rev.*, 1925, v. 26, 44–55.
21. Tolliver D.E., Kyrala G.A and Wing W. H. Observation of the infrared spectrum of helium-hydride molecular ion $^4HeH^+$. *Phys. Rev. Let.*, 1979, v. 43, no. 23, 1719–1722.
22. Crofton M.W., Altman R.S., Haese N.N. and Oka T. Infrared spectra of $^4HeH^+$, $^4HeD^+$, $^3HeH^+$, and $^3HeD^+$. *J. Chem. Phys.*, 1989, v. 91, 5882–5886.
23. Roberge W. and Dalgarno A. The formation and destruction of HeH^+ in astrophysical plasmas. *Astrophys. J.*, 1982, v. 255, 489–496.
24. Engel E.A., Doss N., Harris G.J. and Tennyson J. Calculated spectra for HeH^+ and its effect on the opacity of cool metal-poor stars. *Mon. Not. Roy. Astron. Soc.*, 2005, v. 357, 471–477.
25. Galli D. and Palla F. The chemistry of the early universe. *Astron. Astrophys.*, 1998, v. 335, 403–420.
26. Ketterle W., Figger H. and Walther H. Emission spectra of bound helium hydride. *Phys. Rev. Lett.*, 1985, v. 55, no. 27, 2941–2944.
27. www.nobelprize.org/nobel_prizes/physics/laureates/2001/ketterle.html
28. Ketterle W., Figger H. and Walther H. Emission spectra of bound helium hydride. *Phys. Rev. Lett*, 1985, v. 55, 2941–2944.
29. Ketterle W., Dodhy A. and Walther H. Bound-free emission of the helium hydride molecule. *Chem. Phys. Let.*, 1986, v. 129, no. 1, 76–78.
30. Zirin H. The Solar Atmosphere. Blaisdell Publishing Company, Waltham, MA, 1966.

LETTERS TO PROGRESS IN PHYSICS

The Liquid Metallic Hydrogen Model of the Sun and the Solar Atmosphere VII. Further Insights into the Chromosphere and Corona

Pierre-Marie Robitaille

Department of Radiology, The Ohio State University, 395 W. 12th Ave, Columbus, Ohio 43210, USA.
robitaille.1@osu.edu

In the liquid metallic hydrogen model of the Sun, the chromosphere is responsible for the capture of atomic hydrogen in the solar atmosphere and its eventual re-entry onto the photospheric surface (P.M. Robitaille. The Liquid Metallic Hydrogen Model of the Sun and the Solar Atmosphere IV. On the Nature of the Chromosphere. *Prog. Phys.*, 2013, v. 3, L15–L21). As for the corona, it represents a diffuse region containing both gaseous plasma and condensed matter with elevated electron affinity (P.M. Robitaille. The Liquid Metallic Hydrogen Model of the Sun and the Solar Atmosphere V. On the Nature of the Corona. *Prog. Phys.*, 2013, v. 3, L22–L25). Metallic hydrogen in the corona is thought to enable the continual harvest of electrons from the outer reaches of the Sun, thereby preserving the neutrality of the solar body. The rigid rotation of the corona is offered as the thirty-third line of evidence that the Sun is comprised of condensed matter. Within the context of the gaseous models of the Sun, a 100 km thick transition zone has been hypothesized to exist wherein temperatures increase dramatically from 10^4–10^6 K. Such extreme transitional temperatures are not reasonable given the trivial physical scale of the proposed transition zone, a region adopted to account for the ultra-violet emission lines of ions such as C IV, O IV, and Si IV. In this work, it will be argued that the transition zone does not exist. Rather, the intermediate ionization states observed in the solar atmosphere should be viewed as the result of the simultaneous transfer of protons and electrons onto condensed hydrogen structures, CHS. Line emissions from ions such as C IV, O IV, and Si IV are likely to be the result of condensation reactions, manifesting the involvement of species such as CH_4, SiH_4, H_3O^+ in the synthesis of CHS in the chromosphere. In addition, given the presence of a true solar surface at the level of the photosphere in the liquid metallic hydrogen model, it follows that the great physical extent of the chromosphere is supported by gas pressure, much like the atmosphere of the Earth. This constitutes the thirty-fourth line of evidence that the Sun is comprised of condensed matter.

In order to explain the occurrence of the dark lines in the solar spectrum, we must assume that the solar atmosphere incloses a luminous nucleus, producing a continuous spectrum, the brightness of which exceeds a certain limit. The most probable supposition which can be made respecting the Sun's constitution is, that it consists of a solid or liquid nucleus, heated to a temperature of the brightest whiteness, surrounded by an atmosphere of somewhat lower temperature.

Gustav Robert Kirchhoff, 1862 [1]

1 Introduction

If our current understanding of the solar atmosphere appears strained, it is because the gaseous models of the Sun offer no means, other than elevated temperatures, to account for the presence of highly ionized ions in the corona [2]. As a consequence, temperature values ranging from 10^7–10^{11} K have been inferred to exist in the solar atmosphere [3, p. 172]. Such extreme temperatures should have suggested long ago that the methods utilized to infer coronal temperatures could not be valid, given that the core of the Sun is believed to sustain temperatures of only $\sim 1.6\times10^7$ K [4, p. 9]. The claim that temperatures in localized regions of the corona can be 1 000 times higher than within the solar core, challenges reason.

Furthermore, by accepting elevated coronal temperatures, proponents of the gaseous models must discount the continuous emission of the K-corona as illusionary and produced by the photosphere (see [2] for a completed discussion). The continuous spectrum of the K-corona, devoid of Fraunhofer lines, does closely replicate the emission of the photosphere itself, but the spectrum reddens with elevation [2]. If this spectrum was considered as generated by the corona, then the apparent temperature of the outer solar atmosphere would be no higher than that observed on the surface of the Sun.*

*Note that the apparent temperature of the photosphere (~6 000 K), does not manifest the true energy content of this region. Rather, the author has claimed that it reflects that amount of energy which is contained within the

Should it be true that coronal apparent temperatures are no greater than photospheric values, then it is impossible, within the context of a gaseous Sun, to account for the presence of highly ionized ions (e.g. CaXVII and FeXXIV [6, p. 19]) in the corona. Devoid of condensed matter, the only possible means of generating such ions must rest on temperature. As a result, despite the realization that the spectrum of the K-corona implies that the corona is self-luminous and displays an apparent temperature no higher than that of the photosphere [2], advocates of the gaseous models of the Sun have no choice but to postulate that coronal apparent temperatures far exceed those of the solar surface.

Two problems come to the forefront relative to using elevated temperatures to explain the presence of highly ionized species within the corona. First, extreme temperatures (10^7–10^{11} K [3, p. 172]) must be assumed. Second, the continuous spectrum of the K-corona must be discounted as a byproduct of photospheric light which has been scattered in the solar atmosphere by relativistic electrons (see [2] for a complete discussion).

Moreover, in order to account for the emission lines from ions such as C IV, O IV, and Si IV, gaseous models must incorporate an extremely thin transition zone, whereby apparent temperatures rapidly rise from chromospheric to coronal values over the span of 100 km or less, as illustrated in Fig. 1.

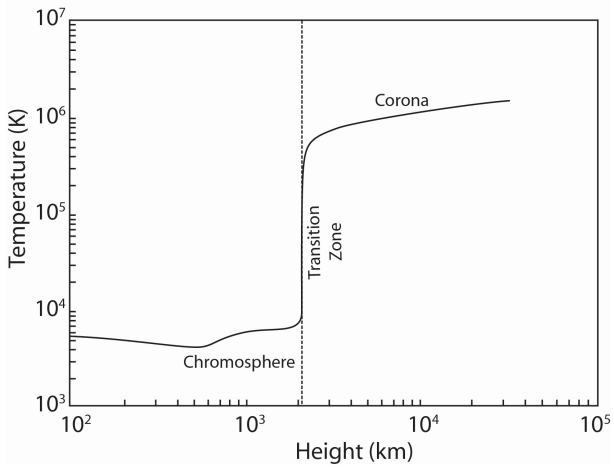

Fig. 1: Schematic representation of the temperature stratification in the solar atmosphere displaying the pronounced increase in the transition zone located at an elevation of ∼2 000 km (dashed line). This figure is based on a discussion presented by Phillips, Feldman, and Landi [6] and is an adaptation of their Fig. 1.1.

2 Temperature Stratification

In his chapter on the chromosphere and corona, P. Ulmschneider states, *"While the corona extends to many solar radii the chromosphere is a layer of only 2 or 3 thousand km thickness which becomes visible near the start and end of a total eclipse. The chromosphere got its name from the prominent red emission of the Hα line of neutral hydrogen at 6563Å. The chromosphere is a layer where the temperature rises from photospheric values of between 4 000 and 6 000 K to about 20 000 K and where neutral hydrogen is still present. In the region of a few 100 km thickness between the chromosphere and corona, called transition layer, hydrogen becomes ionized and the temperature increases from 20 000 to millions of K"* [7, p. 232–233].

A. Bhatnagar outlines that *"Between the upper layer of the chromosphere and corona (although the demarcation is not sharp) lies the 'transition layer', where the temperature rises very steeply, from about 25 000 to 500 000 K in height difference of just 1 000 km"* expanding the extent of the transition region by a factor of 10 [8, p. 32]. Conversely, Phillips, Felman, and Landi emphasize *"Model calculations indeed suggest that the transition zone is extremely thin, less than 100 km"* [6, p. 220].

Such dimensions on the Sun are essentially beyond the limit of reliable detection with current instrumentation. Thus, it is interesting to highlight that *"A growing corpus of observations, particularly those starting with the Skylab mission, showed that the transition zone has a much larger extent than was indicated in the earlier models, leading to a revision of our ideas of its nature..."* [6, p. 210].

Harold Zirin, in candid fashion, reminds his readers that anyone with a ruler can establish that the chromosphere can attain elevations of at least 5 000 km from Hα emissions [9]. He reports that, when viewed in Hα, macrospicules can be seen to extend to 20 000 km [9]. How can neutral hydrogen be found at these heights, if the corona already reaches temperatures of 10^6 K just after the transition zone? If the corona was at millions of degrees, neutral hydrogen should not be found at 20 000 km, a region well within the coronal domain.

The situation is aggravated by the realization that H_{Lym} lines have been known to exist in the corona beyond $1.5R_\odot$ for more than fourty years [10]. This region extends beyond the entire vertical range displayed in Fig. 1. Furthermore, Dermendjiev et al. report, from direct photographic visualization in Hα, that faint lines from neutral hydrogen can be observed far into the corona, causing the authors to postulate how the corona could be 'cooled' to allow for the presence of such a line [11]. Yet, models of the solar atmosphere predict that neutral hydrogen should be absent at elevations beyond 2 000 km, where temperatures approaching 10 000 K already result in the complete ionization of this element (see e.g. Table 4.6 in [12, p. 146–147]).

At the same time, highly ionized Fe lines (FeX–FeXIV) have been used to image the solar corona in great detail and indicate that these species can be found at elevations well inferior to the known locations of neutral hydrogen emission lines [13–16]. Clearly, it is not possible for emission lines

vibrational degrees of freedom found in the photospheric lattice [5].

which, according to the gaseous models of the Sun, require millions of degrees for formation (FeX–FeXIV) to be juxtaposed with Hα lines which are unable to withstand such temperatures. The only solution rests in recognizing that the formation of highly ionized emission lines in the corona stems not from extreme temperatures, but from electron affinity [2]. It should not be inferred that the outer atmosphere of the Sun maintains a temperature stratification which *increases* with increased distance away from the solar body.

3 On The Validity of Temperature Measurements

In order to support the gaseous models, coronal temperatures have been estimated using four key methods [17, p. 178–185]: 1) doppler broadening of emission lines, 2) density gradients in the corona, 3) radio brightness, and 4) ionization equilibrium. All of these methods provide slightly differing answers [3, p. 165–166], but they share a common overarching result: coronal temperatures are thought to be extremely high. In the end, careful analysis reveals that each of these methods is problematic.

3.1 Doppler Broadening of Emission Lines

Doppler broadening of emission lines (e.g. [6, p. 41–43], [17, p. 178–180], [18, p. 90–94]) has been used extensively to set coronal temperatures. The broadening of an emission line, in this case, is assumed to be thermal in nature. The problems of assigning temperatures with such methods are numerous. Zirin [17, p. 178-180] outlines how separate elements can easily produce differing line widths and associated temperatures. Nonetheless, he concludes that valid coronal temperatures can be derived from such methods.

More than fifty years ago, Jefferies and Orrall addressed the problem of obtaining prominence temperatures by employing spectral line widths stating, *"If the broadening is supposed due to thermal motions of the emitting atoms, then, to the extent that the profiles are Gaussian, the hydrogen line widths imply temperatures of over a hundred thousand degrees and the metals of over five million degrees"* [19]. However, it is not possible to have neutral hydrogen present at a temperature of over a hundred thousand degrees, given that the element has been modeled as fully ionized at ~10 000 K (see e.g. Table 4.6 in [12, p. 146–147]).

Jefferies and Orrall continue, *"To avoid the necessity of considering such unacceptably high and discordant temperatures, the hypothesis is frequently made that the line broadening is due both to thermal motions of atoms and to mass motions of small prominence elements having a Maxwellian distribution of velocities. One may, on this basis, compare the widths of lines from two ions of very different atomic weights to find a hypothetical "temperature," T_H and "mean velocity," ξ_H. If the hypothesis is wrong, T_H and ξ_H will, in general, bear no obvious relationship to the kinetic temperature or mean random-velocity fields which they are intended to describe. While the truth of the hypothesis has come more and more to be taken for granted, it seems to us that the evidence in its favor is rather slight and certainly insufficient to allow its uncritical acceptance. We have already . . . suggested that the hypothesis may be invalid for analyzing widths of hydrogen and helium lines in quiescent prominences; in this paper we present evidence for its possible failure in active flare-type events"* [19].

Though the discussion by Jefferies and Orrall cannot be cited in its entirety, the authors go on to make the point that the use of line width analysis could, in fact, lead to *negative* temperatures. Furthermore, they clearly discount the existence of temperatures in the 500 000–1 000 000 K range [19].

Despite Jefferies and Orrall [19], today it is commonplace to infer temperatures from line widths and ascribe any *excessive line shape distortion* to velocity. That is, if the line shape is distorted, either in the low (red) or high (blue) frequency range, net velocities will be added (e.g. see Eq. 2.30 in [6]) which can help account for the distortion. Examples can be found throughout the astrophysical literature (e.g. [20]).

The situation is complicated by the realization that, in addition to thermal effects, the line widths of atoms can be altered by pressure, Stark, and electron broadening mechanisms [21, p. 202-233]. However, the derivation of temperatures from line widths in the solar atmosphere is much more precarious than these considerations or the discussions from Jefferies and Orrall [19] might suggest.

Collisional line broadening with condensed matter could greatly impact the line widths under observation. Such line broadening will be affected by the abundances of condensed material and gaseous atoms in the corona and, most importantly, by the extent of the interaction between any given element and such objects. Furthermore, tight coupling between gaseous atoms and condensed matter could dramatically alter line shapes outside the effects of velocity. In light of the evidence for the presence of condensed matter in the corona [2], all temperature measurements from line widths should be reconsidered.

3.2 Density Gradients

Density gradient approaches rely on the use of the white-light continuous spectrum observed in the corona [17, p. 178–180] or chromosphere [22, p. 170–228]. Modern theory assumes that this spectrum has been produced by scattering photospheric light through the action of relativistic electrons, thereby enabling a temperature for the corona to be inferred [17, p. 111–121]. The difficulty with such an approach lies in the assumption that the corona is not self-luminous and that its spectrum arises from photospheric light which must be scattered. However, if the corona is indeed self-luminous and cool [2], as implied by the presence of neutral hydrogen even up to $1.5R_\odot$ [10], then this entire line of reasoning must be re-evaluated.

3.3 Radio Measurements

Of the four methods for determining coronal temperatures, the final two are perhaps the weakest [17, p. 178–180]. In the end, radio measurements [18, p. 242–247] should be considered with great caution, even though Professor Zirin has stated that they are *"the most dependable data we have"* [9]. Radio data are highly dependent on the input variables (i.e. electron and ion density) which must be modeled in order to obtain brightness temperatures (e.g. see Table I in [9], [12, p. 133–141], and [23]). All determinations of solar brightness temperatures are inherently linked to *a priori* knowledge of electron densities [22, p. 265] which can only be estimated using modeling, *"... it is evident that the quantities $N_e(h)$ and $T_e(h)$ are too inextricably mixed to be seperately derivable from radio observations alone"* [12, p. 137]. Since radio models cannot disentangle electron density from brightness temperatures, they are often guided by results obtained using optical density gradient methods [22, p. 266]. Direct measurements of electron density remain unavailable and theoretical values may not be accurate.

Radio measurements of brightness temperatures are also highly dependent on wavelength and scattering processes (see e.g. [12, p. 133–141], [22, p. 261–271], and [23]). Widely conflicting data can be obtained (e.g. temperatures of only 300 000 at $1.6R_\odot$ [23]). In fact, radio observations appear to be the source of the most extreme temperature values 10^8–10^{10} K [17, p. 128], while scientists remain confronted with addressing values as low as 10^4 K obtained with such methods (see e.g. [12, p. 133–141] and [23]). As a result, it would be imprudent to place an emphasis on coronal or chromospheric temperatures obtained from radio measurements.

3.4 Ionization Equilibrium

It has already been established that ionization calculations result in models of the solar atmosphere which greatly underestimates the presence of neutral hydrogen in the corona. Consequently, it is evident that temperatures derived from ionization equilibrium must be regarded with caution.

As a rule, coronal temperatures derived from ionization equilibrium tend to be too low to accomodate the gaseous models of the Sun [17, p. 181]: *"We must admit, however, that the ionization theory not only gives the wrong temperature, but fails to account for the many stages of ionization observed in the corona. It is possible that temperature variations explain that fact; we can only wait for better observations of the line profiles of intermediate ions to confirm the existence of temperature differences. It is more likely that there is something erroneous in our basic concept of how ionization takes place; but so far, we do not know what this is"* [17, p. 183]. Immediately after writing these lines, Professor Zirin offers what he believes to be the answer: recombination, a process whereby a single electron is captured by an ion leaving it in double excited state, could be much more important in the corona, resulting in a calculated temperature near 2 MK [17, p.184].

In 1966, Zirin had hoped that more UV data would soon be available to lift the cloud of mystery which surrounded ionization equilibrium calculations [17, p. 181–185]. In fact, the new data only added further confusion. Thirty years later, he would write, *"One would think that observations of the solar ultraviolet would solve many of the problems. However, the intensity of these lines was very much lower than expected and to this day images with adequate resolution have not been obtained. While the UV mimics the radio images, brightening in the network, it is impossible to tell if it comes from the spicules or the magnetic regions at their base. The lines show a deep minimum in intermediate ionization stages of C, N, and O ... and the brightness temperature in the extreme ultraviolet scarcely exceeds 4 000 degrees. This gives a remarkable contradiction. Lines are observed of high ionization stages such as carbon 4, neon 5, oxygen 5, which are only formed at temperatures of 100 000 degrees or more but with brightness temperatures 20 times less"* [9]. Nearly forty years after Professor Zirin produced his classic text [17], coronal temperatures from ionization equilibrium are still viewed as too low [3, p. 165–166].

The proper discussion of ionization equilibrium is best reserved for a full treatment. However, suffice to state that methods which depend on the ionization equations are complex (see [24] for a partial review), involving knowledge of whether or not the region of interest can be considered to be in local thermal equilibrium (LTE). E.A. Milne highlighted that the exterior regions of the Sun cannot be considered to adhere to LTE conditions [25, p. 81–83]. Even chromospheric ionization processes depend on non-equilibrium treatments [18, p. 194–198], even if LTE methods continue to be used (see [26] for a brief discussion). Unfortunately, the fluxes associated with such processes remain largely unknown and numerous assumptions will be involved in extracting temperatures with such methods.

In the end, none of the methods utilized to extract coronal temperatures are reliable. Rather, any perceived agreement between approaches is likely to be the result of the desire to *set a reasonable temperature for the corona*. Each method contains enough latitude to permit conformity by altering the value of those input parameters which can only be obtained from theory.

4 The Corona Revisited

Professor Harold Zirin had suspected that *"... there is something erroneous in our basic concept of how ionization takes place"* [17, p. 183]. However, given the belief that the Sun was a gas, no other plausible mechanism of formation could be advanced. Today, the situation has changed dramatically, as a great deal of evidence is building that the Sun is condensed matter (see [2, 24, 27–30] and references therein).

For instance, it is now understood that the corona possesses *"... a radially rigid rotation of 27.5 days synodic period from 2.5 $R_\odot$ to >15$R_\odot$"* [3, p. 116]. This finding by Lewis et al. [31] provides the thirty-third line of evidence that the Sun is comprised of condensed matter. The rigid rotation of the corona is highly suggestive that it possesses condensed matter whose associated magnetic field lines are anchored at the level of the photosphere. Such a structure, if endowed with a elevated electron affinity [2], would provide an elegant network for channeling electrons from the outer reaches of the solar atmosphere onto the photospheric surface. Thus, the corona should be viewed as being in direct contact with the photosphere.

In order to understand ionization states it is important to recognize that condensed matter controls the behavior of the Sun. As previously stated [2], within the solar atmosphere, atoms and ions are being stripped of their electrons by metallic hydrogen present in the corona. Such a process can help ensure that the solar body remains electrically neutral, as electrons are continually conducted back onto the solar surface from the far reaches of the corona. It is known that the electrical conductivity of the corona is extremely high [3, p. 174]. This is in accord with a condensed solar state, which extends into the corona, even if gases are also present in this region.

5 The Chromosphere Revisited

The author has already addressed the chromosphere in detail, as a region of hydrogen re-condensation, superimposed on the corona in the lower portion of the solar atmosphere [28, 29]. He has suggested, that unlike the corona, the chromosphere is not composed of hydrogen in the metallic state. Rather, in the chromosphere, atomic and ionic hydrogen is interacting with other atoms to form hydrides [28, 29] which can be used to build condensed hydrogen structures (CHS). CHS can then bring the harvested hydrogen back onto the solar surface, perhaps using intergrannular lanes [28]. As such, the chromosphere overlaps with the corona. The two regions contain different types of material: metallic in the corona [2] and non-metallic in the chromosphere [28, 29]. Chromospheric material will regain metallic properties once it enters the solar interior, where increased pressures can be used to re-synthesize metallic hydrogen [30].

The tremendous height, 5 000 to 10 000 km, of the chromosphere has posed a longstanding problem for the gaseous models of the Sun [3, p, 140-142]. Early chromospheric models inferred a density scale height of only 150 km [3, p. 140-142]. McCrea [32] attempted to build additional scale height by suggesting that turbulent motions might provide additional support for the chromosphere [3, p. 140-142]. Modern models have extended the theoretical treatment of the scale height problem (see [26] for a brief discussion). But, still today, it remains difficult for the gaseous models of the Sun to account for the presence and extent of the chromosphere. Zirin highlights, *"It was clear that the apparent scale height of 1 000 km far exceeded that in hydrostatic equilibrium. In modern times a convenient solution has been found – denial ... We cannot explain the great height or the erroneous models ... While models place this at 2 000 km, the data say 5 000"* [9]. If it is impossible for the gaseous models to properly account for the great height of the chromosphere, the cause is simple to understand. It is not possible for a gas to support itself. But relative to structural support, gas pressure has been utilized in modern solar theory to explain why a gaseous Sun does not collapse on itself. However, such arguments have been discounted, precisely because a gaseous object cannot possess true surfaces [33]. Without a support mechanism, a gaseous Sun cannot exist [33].

Conversely, within the context of a condensed solar body [33], the Sun does not collapse upon itself because liquids and solids are essentially incompressible. Furthermore, unlike the case with the gaseous Sun, the chromosphere can now be easily supported using gas pressure. This same mechanism is responsible for the support of the Earth's atmosphere (see [33] for a larger discussion). When a gaseous atom encounters a real surface, it reverses its course creating a net upward force. Such a mechanism provides a genuine means of supporting the chromosphere and thereby constitutes the thirty-fourth line of evidence that the Sun is condensed matter (see [2, 24, 27–30, 33] and references therein for the others).

6 The Transition Zone Revisited

Within the gaseous models of the Sun, a transition zone has been conceived in order to account for the existence of ions with *intermediate* levels of ionization. Species such as C IV, O IV, and Si IV come to mind in this regard. Since the intensity of all transition zone lines are low, modern models simply create an extremely narrow region of the solar atmosphere to account for this lack of signal, as illustrated in Fig. 1. Nonetheless, C IV, O IV, and Si IV remain interesting, as they could be created by stripping hydrides such as CH_4, H_3O^+, and SiH_4 of their hydrogen [28]. The vibrational signatures of these molecules (the C-H, O-H, and Si-H stretches) have been observed on the Sun [34]. The author has already suggested that the chromosphere is a region of hydrogen re-condensation where hydrides play an important role [28, 29]. It remains reasonable to conclude that the transition zone does not exist. Rather, the ions which are currently associated with this region of the solar atmosphere are simply involved in the transfer of multiple protons and electrons onto the condensed hydrogen structures, CHS, which constitute the chromosphere. This region of the solar atmosphere therefore plays a vital role in preserving the mass of the Sun and ensuring that metallic hydrogen can eventually be re-synthesized within its interior.

7 Conclusion

Through a recent series of publications (most notably [2, 28, 29]), the author has endeavored to alter our understanding of the solar atmosphere. Rather than a chaotic assembly of gaseous plasma, the chromosphere and corona become the site of both structure and function in the Sun. Such structure is dismissed by the gaseous models, whose advocates prefer to speak of visualizing *"force balance"* [26], rather than real objects. At the same time, the history observational solar physics is replete with scientists, like Father Angelo Secchi, who believed that they were seeing real structures on the Sun [35, 36]. In a parallel line of reasoning, the gaseous models provide no true function, either for the chromosphere or the corona. Conversely, in the liquid metallic model, the corona harnesses electrons [2], the chromosphere condenses hydrogen atoms [28, 29]. In the corona, highly ionized ions are produced when their parent atoms, or ions, come into contact with metallic hydrogen which possesses an elevated electron affinity. They are thereby stripped of their electrons [2]. The metallic hydrogen which is present in the corona has been projected into the solar atmosphere from its site of formation below the surface of the Sun [2]. Since condensed matter appears likely to exist in the corona, it is not tremendously hot, but maintains an apparent temperature which decreases with elevation from the solar body. In the chromosphere, where non-metallic condensed hydrogen structures are formed, the ionization states revealled from emission lines are linked to key hydride based chemical processes [28, 29]. The transition zone does not exist. It serves a purpose only in the context of the gaseous solar models. Much has been advanced recently relative to the condensed nature of the Sun [2, 24, 27–30, 33] and much remains to be considered. In the end, given the ever mounting evidence for condensed matter (see [2, 24, 27–30, 33] and references therein), eventually the elegance and simplicity of these models will surely come to be recognized.

Acknowledgment

Luc Robitaille is acknowledged for figure preparation.

Dedication

This work is dedicated to the memory of Captain Corona [37], Professor Harold Zirin, whose books [17,18] and articles (e.g. [9]) were both illuminating in their discourse and refreshing in their candor.

Submitted on: May 30, 2013 / Accepted on: May 30, 2013
First published online on: May 31, 2013

References

1. Kirchhoff G. The physical constitution of the Sun. In: *Researches on the Solar Spectrum and the Spectra of the Chemical Elements*. Translated by H.E. Roscoe, Macmillan and Co., Cambridge, 1862, p. 23.
2. Robitaille P.M. The Liquid Metallic Hydrogen Model of the Sun and the Solar Atmosphere V. On the nature of the corona. *Progr. Phys.*, 2013, v. 3, L22-L25.
3. Bhatnagar A. and Livingston W. Fundamentals of Solar Astronomy (World Scientific Series in Astronomy and Astrophysics - Vol. 6), World Scientific, New Jersey, 2005.
4. Kippenhahn R. and Weigert A. Stellar structure and evolution. Springer-Verlag, Berlin, 1990.
5. Robitaille P.M. On the temperature of the photosphere: Energy partition in the Sun. *Progr. Phys.*, 2011, v. 3, 89–92.
6. Phillips K.J.H., Feldman U. and Landi E. Ultraviolet and X-ray Spectroscopy of the Solar Atmosphere. Cambridge University Press, Cambridge, 2008.
7. Ulmschneider P. The physics of the chromospheres and coronae. In: *Lectures on Solar Physics* (H.M. Antia, A. Bhatnagar and R. Ulmschneider, Eds.), Springer, Berlin, 2003, p. 232–280.
8. Bhatnagar A. Instrumentation and observational techniques in solar astronomy. In: *Lectures on Solar Physics* (H.M. Antia, A. Bhatnagar and R. Ulmschneider, Eds.), Springer, Berlin, 2003, p. 27–79.
9. Zirin H. The mystery of the chromosphere. *Solar Phys.*, 1996, v. 169, 313–326.
10. Gabriel A.H., Garton W.R.S., Goldberg L., Jones T.J.L., Jordan C., Morgan F.J., Nicholls R.W., Parkinson W.J., Paxton H.J.B., Reeves E.M., Shenton C.B., Speer R.J. and Wilson R. Rocket observations of the ultraviolet solar spectrum during the total eclipse of 1970 March 7. *Astrophys. J.*, 1971, v. 169, 595–614.
11. Dermendjiev V.N, Mouradian Z., Duchlev P. and Leroy J.L. Faint H alpha emission in the solar corona: Morphological, situational and hydromagnetic analysis. *Solar Phys.*, 1994, v. 149, no. 2, 267–277.
12. Bray R.J. and Loughhead R.E. The Solar Chromosphere (The International Astrophysics Series), Chapman and Hall, London, U.K., 1974.
13. Wood B.E., Karovska M., Cook J.W., Brueckner G.E., Howard R.A., Korendyke C.M. and Socker D.G. Search for brightness variations in FeXIV coronagraph observations of the quiescent solar corona. *Astrophys. J.*, 1998, v. 505, 432–442.
14. Habbal S.R., Druckmüller M., Morgan H., Daw A., Johnson J., Ding A., Arndt M., Esser R., Rušin V. and Scholl I. Mapping the distribution of electron temperature and Fe charge states in the corona with total solar eclipse observations. *Astrophys. J.*, 2010, v. 708, 1650–1662.
15. Habbal S.R., Druckmüller M., Morgan H., Scholl I., Rušin V., Daw A., Johnson J. and Arndt M. Total solar eclipse observations of hot prominence shrouds. *Astrophys. J.*, 2010, v. 719, 1362–1369.
16. Habbal S.R., Morgan H. and Druckmüller M. A new view of coronal structures: Implications for the source and acceleration of the solar wind – First Asia-Pacific Solar Physics Meeting. *ASI Conf. Ser.*, 2011, v. 2, 259–269.
17. Zirin H. The Solar Atmosphere. Blaisdell Publishing Company, Waltham, M.A., 1966.
18. Zirin H. Astrophysics of the Sun. Cambridge University Press, Cambridge, U.K., 1988.
19. Jefferies J.T and Orrall F.Q. On the interpretation of prominence spectra. II. The line and continuous spectrum of the spray-type limb event of March 7, 1959. *Astrophys. J.*, 1961, v. 133, 946–962.
20. Fleck B., Brekke P., Haugan S., Sanchez Duarter L., Domingo V., Gurman J.B. and Poland A.I. Four years of SOHO discoveries – Some highlights. *ESA Bulletin 102*, May 2000. (Available online sohowww.nascom.nasa.gov/publications/ESA_Bull102.pdf).
21. Gray D.F. The observation and analysis of stellar photospheres (Cambridge Astrophysics Series), 2nd Edition, Cambridge University Press, Cambridge, U.K., 1995.
22. Thomas R.N. and Athay R.G. Physics of the Solar Chromosphere. Interscience Publishers, New York, N.Y., 1961.

23. Thejappa G. and MacDowall R.J. Effects of scattering on radio emission from the quiet Sun at low frequencies. *Astrophys. J*, 2008, v. 676, no. 2, 1338–1345.
24. Robitaille P.M. Liquid metallic hydrogen II. A critical assessment of current and primordial helium levels in the Sun. *Progr. Phys.*, 2013, v. 2, 35–47.
25. Milne E.A. Thermodynamics of the stars. *Handbuch der Astrophysik*, 1930, v. 3, Part 1, 65–255 (also in Menzel D.H. Selected Papers on the Transfer of Radiation: 6 papers by Arthur Schuster, K. Schwarzschild, A.S. Eddington, S. Rosseland, E.A. Milne. Dover Publications, Inc., New York, 1966, 77–269).
26. Judge P. Observations of the solar chromosphere (Solar MHD Theory and Observations: A High Spatial Resolution Perspective). *ASP Conf. Series*, 2006, v. 354, 259–275.
27. Robitaille P.M. The liquid metallic hydrogen model of the Sun and the solar atmosphere III. Importance of continuous emission spectra in flares, coronal mass ejections, prominences, and other coronal structures. *Progr. Phys.*, 2013, v. 3, L11–L14.
28. Robitaille P.M. The liquid metallic hydrogen model of the Sun and the solar atmosphere IV. On the nature of the chromosphere. *Progr. Phys.*, 2013, v. 3, L15–L21.
29. Robitaille P.M. The liquid metallic hydrogen model of the Sun and the solar atmosphere VI. Helium in the chromosphere. *Progr. Phys.*, 2013, v. 3, L26–L28.
30. Robitaille P.M. Liquid metallic hydrogen: A building block for a liquid Sun. *Progr. Phys.*, 2011, v. 3, 60–74.
31. Lewis D.J., Simnett G.M., Brueckner G.E., Howard R.A., Lamy P.L. and Schwenn R. LASCO observations of the coronal rotation. *Solar Phys.*, 1999, v. 184, no. 2, 297–315.
32. McCrea W.H. The hydrogen chromosphere. *Mon. Not. Roy. Astron. Soc.*, 1929, v. 29, 483–497.
33. Robitaille J.C. and Robitaille P.M. Liquid Metallic Hydrogen III. Intercalation and Lattice Exclusion Versus Gravitational Settling and Their Consequences Relative to Internal Structure, Surface Activity, and Solar Winds in the Sun. *Progr. Phys*., 2013, v. 2, 87-97.
34. Sinha K. Molecules in the Sun. *Proc. Astron. Soc. Australia*, 1991, v. 9, 32–36.
35. Robitaille P.M. A thermodynamic history of the solar constitution – I: The journey to a gaseous Sun. *Progr. Phys.*, 2011, v. 3, 3–25.
36. Robitaille P.M. A thermodynamic history of the solar constitution - II: The theory of a gaseous Sun and Jeans' failed liquid alternative. *Progr. Phys.*, 2011, v. 3, 41–59.
37. Caltech News 1/12/2012, Harold Zirin, 82 (http://www.caltech.edu/article/295 Accessed online on 5/23/2013).

The Journal on Advanced Studies in Theoretical and Experimental Physics, including Related Themes from Mathematics

PROGRESS IN PHYSICS

A quarterly issue scientific journal, registered with the Library of Congress (DC, USA). This journal is peer reviewed and included in the abstracting and indexing coverage of: Mathematical Reviews and MathSciNet (AMS, USA), DOAJ of Lund University (Sweden), Zentralblatt MATH (Germany), Scientific Commons of the University of St. Gallen (Switzerland), Open-J-Gate (India), Referativnyi Zhurnal VINITI (Russia), etc.

Electronic version of this journal:
http://www.ptep-online.com

Postal Address

Department of Mathematics and Science,
University of New Mexico,
705 Gurley Ave., Gallup, NM 87301, USA

ISSN: 1555-5534 (print)
ISSN: 1555-5615 (online)

Standard Address Number: 297-5092
Printed in the United States of America

OCTOBER 2013 **VOLUME 4**

CONTENTS

Information for Authors and Subscribers

Progress in Physics has been created for publications on advanced studies in theoretical and experimental physics, including related themes from mathematics and astronomy. All submitted papers should be professional, in good English, containing a brief review of a problem and obtained results.

All submissions should be designed in LaTeX format using *Progress in Physics* template. This template can be downloaded from *Progress in Physics* home page http://www.ptep-online.com. Abstract and the necessary information about author(s) should be included into the papers. To submit a paper, mail the file(s) to the Editor-in-Chief.

All submitted papers should be as brief as possible. We accept brief papers, no larger than 8 typeset journal pages. Short articles are preferable. Large papers can be considered in exceptional cases to the section *Special Reports* intended for such publications in the journal. Letters related to the publications in the journal or to the events among the science community can be applied to the section *Letters to Progress in Physics*.

All that has been accepted for the online issue of *Progress in Physics* is printed in the paper version of the journal. To order printed issues, contact the Editors.

This journal is non-commercial, academic edition. It is printed from private donations. (Look for the current author fee in the online version of the journal.)

Space-Time Uncertainty and Cosmology: a Proposed Quantum Model of the Universe

Sebastiano Tosto

Italy, E-mail: stosto@inwind.it, stosto44@gmail.com

The paper introduces a cosmological model of the quantum universe. The aim of the model is (i) to identify the possible mechanism that governs the matter/antimatter ratio existing in the universe and concurrently to propose (ii) a reasonable growth mechanism of the universe and (iii) a possible explanation of the dark energy. The concept of time-space uncertainty, on which is based the present quantum approach, has been proven able to bridge quantum mechanics and relativity.

1 Introduction

Physical cosmology is the science of the most fundamental questions about past, present and future of the universe. Born in the modern form with the early Einstein general relativity (1916), it involves today all branches of the theoretical physics. The conceptual basis of cosmology relies not only on the theories of gravity field, but also on the fundamental interactions between elementary particles. Likely the first attempt of extending the achievements of general relativity to propose a model of universe based on a physical theory was made by Einstein himself with the introduction of the cosmological constant Λ. At that time the quantum theory was at its very early beginning, while the gravitational interaction seemed the most general physical law governing the dynamics of celestial bodies; so the relativity, with or without Λ, soon appeared as the most valuable resource to proceed beyond the Newton physics.

The first milestone of the modern cosmology is due to Friedmann (1922) and (1924); the hypothesis of universe homogeneous and isotropic allowed inferring the equations that describe shape and expansion/contraction propensity of the universe depending on the value of the density parameter Ω. After these early contributions, have been proposed several models of universe, e.g. by Lemaitre (1929) and Eddington (1930).

The first experimental milestone of cosmology is due to Hubble, who measured the Doppler shift of light emitted by far galaxies (1929): the experimental data revealed the recession velocity law of galaxies with respect to earth. Since then, any model of universe should allow for this experimental evidence. The second experimental landmark was the discovery of the cosmic microwave background radiation (Penzias and Wilson, 1965).

An essential added value to the theoretical cosmology came from the almost simultaneous development of quantum mechanics. Without this physical background and the recent Standard Model, the modern cosmology would be inconceivable. The cosmic abundance of elements has been investigated by Weizsacker (1938) and then by Gamow et al (1948); Chandrasekhar (1942) and more recently Fowler et al [1] pointed out several processes in the stars that concurrently account for the formation of heavy elements in the universe.

On the one hand, the understanding of the nuclear processes explains the existence of stars and other objects (quasars, white dwarf and so on); on the other hand, however, is the general relativity that explains the existence and features of the black holes. The crucial point of the modern physics and cosmology is the difficulty of merging relativistic and quantum theories. Several papers have been published on quantum gravity, e.g. [2,3]. Today the string theory is deemed to be a step towards the unification of both theories [4,5]; unavoidably the string theory has been also implemented by cosmologists to investigate problems of mere quantum nature, like for instance the vacuum energy and the dark energy [6], and the cosmological constant as well [7,8,9]. However, the mathematical difficulties of these theories are daunting, and their previsions hardly testable.

Yet to shed light on fundamental issues of cosmology are also useful plain models that exploiting simple assumptions allow reliable order of magnitude estimates; simplified models are functional to focus essential but even so significant information.

The present paper aims to infer the order of magnitude estimates starting from a quantum standpoint. The input values implemented in this paper are the literature estimates of the universe diameter $d_u = 8.7\times10^{26}$ m and age $t_u = 4.3\times10^{17}$ s. The total mass of the universe reported in the literature is estimated to be about $m_u = 3\times10^{52}$ kg, counting however the stars only. Thus it is reasonable to expect that the effective value M_u of total mass should actually be considerably greater than m_u. Indeed this latter does not include contributions like the dark mass or the total mass of all black holes possible existing in our universe, which instead should be also taken into account when correlating these three main features of the universe; this reasonably suggests $M_u > m_u$. The fourth key value to be introduced is the expansion rate on the universe, usually expressed through the Hubble constant $H_0 = 2.3\times10^{-18}$ s^{-1}; this number, which presumably averages the value of a true function of time, has been object of great debate because of its importance in cosmology.

2 Quantum background

Physicists believe unsatisfactory a theory based on the wave function ψ without direct physical meaning, e.g. [10]; indeed $\psi^*\psi$ only has the statistical meaning of probability density and contains the maximum information obtainable about a physical system. Moreover also the Wigner function [11], although providing significant information about the quantum states, presents conceptual difficulties: it cannot be really regarded as a probability distribution in the classical sense, it is a quasi-probability that can take negative values; moreover it can represent the average value of an observable but not, in general, also its higher power moments. These difficulties, both inherent the wave formalism, are overcome in a model that exploits directly the statistical formulation of the quantum uncertainty, which becomes itself a fundamental assumption of the model and reads in one space dimension

$$\Delta x\, \Delta p_x = n\hbar = \Delta\varepsilon\, \Delta t. \qquad (2,1)$$

The second equality is formally obtained from the former rewritten $(\Delta x/\mathrm{v}_x)(\mathrm{v}_x \Delta p_x) = n\hbar$ with the same number n of states and defining $\mathrm{v}_x = \Delta x/\Delta t$ and $\Delta\varepsilon = \mathrm{v}_x \Delta p_x$; these definition hold because n and the uncertainty ranges are arbitrary. (2,1) compel the positions

$$x \to \Delta x; \qquad p_x \to \Delta p_x; \qquad t \to \Delta t; \qquad \varepsilon \to \Delta\varepsilon. \qquad (2,2)$$

No further hypothesis is necessary besides that of waiving the random local values of the dynamical variables, considered random, unknown and unpredictable. To clarify the kind of quantum approach required by the positions (2,2) and highlight why (2,1) have prospective interest also in cosmology, are useful two examples shortly sketched below. The quantum properties are inferred implementing directly the physical definitions of the observable of interest, without solving the pertinent wave equations; note however that the operator formalism of wave mechanics is also obtained as a corollary of these equations [12], which explains why the results are anyway the same.

The first example concerns the angular momentum $\mathbf{M} = \mathbf{r} \times \mathbf{p}$ whose component along the arbitrary unit vector $\mathbf{w}$ is $M_w = \mathbf{r} \times \mathbf{p} \cdot \mathbf{w}$; the vectors are defined in a reference system R. The positions (2,2) compel $\mathbf{r} \to \Delta\mathbf{r}$ and $\mathbf{p} \to \Delta\mathbf{p}$ to calculate the number l of states consistent with the ranges $\Delta\mathbf{r}$ and $\Delta\mathbf{p}$ physically allowed to the particle. Thus $M_w = (\Delta\mathbf{r} \times \Delta\mathbf{p}) \cdot \mathbf{w} = (\mathbf{w} \times \Delta\mathbf{r}) \cdot \Delta\mathbf{p}$ yields $M_w = \Delta\mathbf{W} \cdot \Delta\mathbf{p}$, where $\Delta\mathbf{W} = \mathbf{w} \times \Delta\mathbf{r}$. So $M_w = 0$ if $\Delta\mathbf{p}$ and $\Delta\mathbf{W}$ are orthogonal; else, rewriting $\Delta\mathbf{W} \cdot \Delta\mathbf{p} = (\Delta\mathbf{p} \cdot \Delta\mathbf{W}/|\Delta\mathbf{W}|)\,|\Delta\mathbf{W}|$ one finds $\pm\Delta p_W = \Delta\mathbf{p} \cdot \Delta\mathbf{W}/|\Delta\mathbf{W}|$ and thus $M_w = \pm\Delta W \Delta p_W$, i.e. $M_w = \pm l\hbar$ according to (2,1). One component of $\mathbf{M}$ only is knowable; repeating the same approach for another component trivially means changing $\mathbf{w}$. Therefore the average values $< M_x^2 >$, $< M_y^2 >$ and $< M_z^2 >$ calculated in the same way should be equal. The components are averaged over the possible states summing $(l\hbar)^2$ from $-L$ to $+L$, where L is an arbitrary maximum value of l; so $< M_i^2 > = \sum_{l_i=-L}^{l_i=L} (\hbar l)^2/(2L+1)$ i.e. $M^2 = \sum_{i=1}^{3} < M_i^2 > = L(L+1)\hbar^2$. The mere physical definition of angular momentum is enough to find quantum results completely analogous to that of the wave mechanics without any hypothesis on the angular motion. The same holds for the energy levels of hydrogenlike atoms. The concerned definitions are now the energy $\varepsilon = p^2/2m - Ze^2/r$, being m the electron mass, and the momentum $p^2 = p_r^2 + M^2/r^2$. The positions (2,2) $p_r \to \Delta p_r$ and $r \to \Delta r$ yield $\Delta\varepsilon = \Delta p_r^2/2m + M^2/2m\Delta r^2 - Ze^2/\Delta r$. Two numbers of states are expected because of the radial and angular uncertainties. The positions (2,2) and the previous result yield $\Delta\varepsilon = n^2\hbar^2/2m\Delta r^2 + l(l+1)\hbar^2/2m\Delta r^2 - Ze^2/\Delta r$ that reads also $\Delta\varepsilon = \varepsilon_o + l(l+1)\hbar^2/2m\Delta r^2 - E_{el}$ with $E_{el} = Z^2e^4m/2n\hbar^2$ and $\varepsilon_o = (n\hbar/\Delta r - Ze^2m/n\hbar)^2/2m$. Minimizing $\Delta\varepsilon$ with $\varepsilon_o = 0$ yields $\Delta r = n^2\hbar^2/Ze^2m$; so $l \le n-1$ in order to get $\varepsilon < 0$, i.e. a bound state; $\varepsilon_{rot} = l(l+1)E_o/n^4$ yields the rotational energy of the atom as a whole. Also here appears that the range sizes do not play any role in determining the energy levels. The physical meaning of Δr, the early Bohr radius, appears noting that actually $E_{el} = -Ze^2/2\Delta r$, i.e. E_{el} is the energy of two charges of opposite sign delocalized within a diametric distance $2\Delta r$ apart. It appears now that the quantum numbers of the eigenvalues are actually numbers of allowed states of quantum systems.

The key point of this introduction is not the chance of having found well known results, but the fact of having extended this kind of approach to the special and general relativity [13,14]; selected results of interest for the purposes of the present paper are reported in the appendix. In this respect, some relevant features of this approach will be exploited later and thus deserve attention.

- Both time and space coordinates are by definition inherent any model based on (2,1).
- Any uncertainty range is defined by two boundary values, e.g. $\Delta x = x_1 - x_0$; either of them is necessarily defined with respect to the origin of a reference system, the other one controls the range size. Since both x_0 and x_1 are arbitrary, unknown and unknowable by assumption, neither size nor reference system are specified or specifiable. Any result obtained from $\mathbf{M} = \mathbf{r} \times \mathbf{p}$ depends on the particular R where are defined $\mathbf{r}$ and $\mathbf{p}$. Yet, once having introduced the positions (2,2), any reference to the initial R is lost, whereas the eigenvalues are correctly inferred from $\Delta\mathbf{r}$ and $\Delta\mathbf{p}$ only; indeed $\Delta\mathbf{M} = \Delta\mathbf{r} \times \Delta\mathbf{p}$ yields a range $\Delta\mathbf{M}$ of angular momenta corresponding to all values of the arbitrary number n of states concurrently introduced via (2,1). Otherwise stated, the previous examples have shown that the boundary values r_0 and r_1 of each i-th component Δr_i are unnecessary and do not play any role to find the eigenvalues; so, since the same holds also for the momentum range, once disregarding both coordinates

neither the range sizes nor the reference system are in fact specifiable. Hence, in general, privileged reference systems are inherently excluded by the agnostic form of space-time uncertainty of (2,1), i.e. the results hold in any four dimensional reference system.

- These examples emphasize that both boundary coordinates could even be time dependent without changing approach or result: once ignoring the local dynamical variables, conceptually and not to simplify or approximate some calculation, no information on the ranges is actually required.
- The positions (2,2) skip the necessity of solving the pertinent wave equations and allow working directly on the physical definitions of the observables; (2,1) extract the allowed quantum information from the analytical form itself of the equation defining the observable.
- The concept of delocalization resulting from (2,1) has more agnostic meaning than that of the wave formalism: here is waived even the concept of probability density.
- (2,1) and the positions (2,2) rule out the classical concept of distance, because the local coordinates that define the distance are disregarded themselves "a priori"; this means that comoving and proper distances cannot in fact be calculated, while saving however their conceptual physical meaning.

Two questions arise at this point: are (2,1) usefully applicable also in cosmology? If they really do, why not think that even the physical dimensions of G could be regarded like that of the angular momentum previously sketched? Nothing excludes "a priori" positive answers, which however imply clearly that the universe is understandable like a quantum object. In fact is just this the crucial point that justifies the present model. These quantum examples have been shortly introduced to highlight the strategy of the present paper, i.e. to emphasize the role of the space-time quantum uncertainty in cosmology. The same kind of approach will be extended to the physics of the universe exploiting both (2,1) to implement G via its physical dimensions: the idea is to regard the physical definition of G likewise as done with the angular momentum. Accordingly the gravity constant is not a mere numerical value, but a physical amount defined by its dimensional factors. In effect, at least in principle, nothing prevents regarding the numerical value of G as that resulting from a combination of mass and time and space uncertainties; so these factors can be replaced by the respective time-space ranges that characterize the properties of the universe and handled exactly as done previously. Three examples useful in the following are highlighted below.

Write $G = \Delta r^3 m^{-1} \Delta t^{-2}$ and calculate

$$\delta G = (dG/d\Delta r)_0 \, \delta\Delta r + (dG/d\Delta t)_0 \, \delta\Delta t + (dG/dm)_0 \, \delta m$$

in an arbitrary reference system R; the subscript emphasizes that the derivatives are calculated at arbitrary Δr_0, m_0 and Δt_0. Apparently a well defined value of gravity constant seems inconsistent with the arbitrariness of Δt, Δr and m inherent its physical dimensions and required by the positions (2,2). Yet the chance of compelling $\delta G = 0$ establishes a constrain on the variability of the constituent factors that makes the definition of G compatible even with a constant value; moreover this constrain is ensured at any age of the universe just because of the arbitrary values of Δr_0 and m_0 that represent its size and total mass at any age Δt_0. So the problem is not the constancy of G, but that of demonstrating a sensible physical meaning of the constrain itself. Divide both sides of the previous expression by $\Delta r_0^3/(m_0 \Delta t_0^2)$ and put $\delta G = 0$; this is not necessarily true because some theories regard G as time dependent function [15, 16], yet let us implement for simplicity this usual position. Here $\delta m \neq 0$ because some models of universe, the so called self-creation cosmology models [17], introduce mass production as a function of time. One finds thus $3\delta\Delta r/\Delta r_0 - \delta m/m_0 - 2\delta\Delta t/\Delta t_0 = 0$. Exploit the fact that the range sizes are arbitrary and that the increments $\delta\Delta r$, δm and $\delta\Delta t$ are arbitrary as well and of course defined independently of Δr_0, m_0 and Δt_0; then regard

$$\left(\frac{3}{2} - \frac{\Delta r_0}{2m_0}\frac{\delta m}{\delta\Delta r}\right)\delta\Delta r = \frac{\Delta r_0}{\Delta t_0}\,\delta\Delta t$$

in order that this equation has in particular a physical meaning of specific interest for the present model. So let us write

$$a(t) = \frac{3}{2} - \frac{\Delta r_0}{2m_0}\frac{\delta m}{\delta\Delta r}; \qquad c = \frac{\Delta r_0}{\Delta t_0}; \qquad \delta\Delta r = \frac{c}{a(t)}\,\delta\Delta t$$

where $a(t)$ is a dimensionless arbitrary function of time. Consider now the particular case of very small range size increments via the positions $\delta\Delta r \to dr$ and $\delta\Delta t \to dt$, possible just because of their arbitrariness, and integrate both sides of the former equation between two arbitrary r_1 and r_2 to which correspond the respective times t_1 and t_2 necessary for a photon to travel the space range $\chi = r_2 - r_1$. Of course the integration reads $\chi = \int_{t_1}^{t_2} a(t)^{-1} c dt$. Therefore with these integration limits and this definition of the constant ratio $\Delta r_0/\Delta t_0$, the resulting equation has the well known physical meaning of particle horizon distance and introduces the concept of scale function $a(t)$.

To complete this analysis on the physical dimensions of G, put $\delta m \to dm$ consistently with dr and dt and consider that the equation of $a(t)$ takes the form $dm = \alpha(3/2 - a(t))dr$, where $\alpha = 2m_0/\Delta r_0$; having defined $dr = cdt/a(t)$, one finds $dm/\alpha = 3ca(t)^{-1}dt/2 - cdt$. The integral of this equation between the fixed times t_1 and t_2 arbitrarily defined and the corresponding m_1 and m_2 yields $(m_2 - m_1)/\alpha = 3(r_2 - r_1)/2 - c(t_2 - t_1)$. In general an equation having the form $\alpha^{-1}\delta m = 3\delta r/2 - c\delta t$ does not have specific physical meaning, because the quantities at right hand side are arbitrary; for instance

$\delta m = 0$ if in particular $\delta r = 2c\delta t/3$, whereas any other value of $\delta m \neq 0$ would be in principle allowed as well. This simply emphasizes that the physical meaning of $a(t)$ is not hampered by constrains on the values of δm or $\delta\Delta t$ or χ. Yet it is also possible to split the equation into $m_2/\alpha - 3r_2/2 + ct_2 = r_0$ and $m_1/\alpha - 3r_1/2 + ct_1 = r_0$, with r_0 arbitrary, which read thus $m_2/\alpha = \delta r_2^*$ and $m_1/\alpha = \delta r_1^*$ with $\delta r_2^* = r_0 + 3r_2/2 - ct_2$ and $\delta r_1^* = r_0 + 3r_1/2 - ct_1$. These equations have in effect a well defined physical meaning, because they read $m_2/\delta r_2^* = m_1/\delta r_1^* = const$. The chance of having inferred from G an equation having the form $m/\delta r^* = const$ is important because it links uniquely any mass m to a corresponding range δr^* via a proportionality factor *const*; as this link must necessarily involve G via a constant term, one expects by dimensional reasons that necessarily $const \propto G/c^2$. Before concerning this point, note that these results have been obtained simply defining $G = \Delta r^3 m^{-1}\Delta t^{-2}$, rather than by implementing additional hypotheses; thus this way of regarding G contains inherently concepts essential to describe an expanding universe.

To better understand the last result, let us consider a further way to exploit the physical dimensions of G via (2,1). Rewrite $G = \Delta r^3/(m\Delta t^2)$ as $\Delta r = Gm/\mathrm{v}^2$ with $\mathrm{v} = \pm\Delta r/\Delta t$; so v is the average velocity necessary for a particle to travel Δr during a time range Δt in any R, as stressed before. The maximum value allowed to v, defined along one coordinate axis for simplicity, introduces a minimum range size Δr_0 of Δr given by $\Delta r_0 = Gm/c^2$. By definition Δr_0 is the distance traveled by a photon starting from an arbitrary point, defined without loss of generality as the origin of R. Since the photon can move around the origin towards the negative or positive side of the reference axis with equal probability, as indeed either sign of v is identically admissible, Δr_0 is one half of a total uncertainty range Δr_s where the photon is certainly enclosed; so $\Delta r_s = 2\Delta r_0$ yields

$$\Delta r_s = 2Gm/c^2 \qquad (2,3)$$

that defines therefore the boundary of the space range outside which the photon cannot escape. This range size has a general physical meaning characterized by the ratio $m/\Delta r_s$ only; also, the same holds of course for a massive particle having $\mathrm{v} < c$. This equation, already inferred in a more general way still via (2,1) only [18], has the same form just found examining $a(t)$: here we simply acknowledge that $const = 2G/c^2$.

Consider eventually that (2,1) read $\Delta x = (\Delta\varepsilon/\Delta p_x)\Delta t$; moreover it is shown in the appendix that $\Delta p_x = \mathrm{v}_x\Delta\varepsilon/c^2$, so that $\Delta x^3 = (c^2/\mathrm{v}_x)^3\Delta t^3$. Dividing both sides of this equation by $m\Delta t^2$ one finds $\Delta x^3/(m\Delta t^2) = (c^2/\mathrm{v}_x)^3\Delta t/m$. Hence

$$\frac{\Delta x^3}{m\Delta t^2} = \frac{c^3}{\xi^3}\frac{\Delta t}{m}; \qquad \mathrm{v_x} = \frac{\mathrm{c}^2\Delta\mathrm{t}}{\Delta\mathrm{x}}; \qquad \xi = \frac{\mathrm{v_x}}{\mathrm{c}}; \qquad \xi < 1. \quad (2,4)$$

Define $\xi = \xi_G\,\xi_c$, so that the right hand side of the first (2,4) reads $(c/\xi_c)^3\Delta t/m$ and the left hand side $\xi_G^3\Delta x^3/(m\Delta t^2)$. Moreover regard in particular $\Delta t \equiv \Delta t_u$ and $\Delta x \equiv \Delta r_u$; this is certainly possible because all range sizes of (2,1) are arbitrary, so they can be regarded with reference to any specific case of interest. It is also possible to define ξ_G in order that the left hand side term corresponds to the value of G with the known values of Δr_u and Δt_u, so that (2,4) yields also the value of ξ_c; in other words (2,4) splits as follows

$$G = \xi_G^3\frac{\Delta r_u^3}{m\Delta t_u^2}; \qquad G = \frac{c^3}{\xi_c^3}\frac{\Delta t_u}{m}; \qquad \xi = \xi_G\,\xi_c < 1. \qquad (2,5)$$

The previous considerations have evidenced that both expressions are compatible with a constant value of G. The problem is to show that in this way ξ effectively verifies the required inequality. The numerical results for $m \equiv m_u$ yield $\xi_G = 0.17$ and $\xi_c = 1.79$, i.e. $\xi = 0.3$. According to (2,4) ξ does not depend directly on m, whereas (2,5) show that ξ_G and ξ_c do. For instance, repeating the calculation with $m \equiv 10m_u$ at the same Δt_u one would find $\xi_G = 0.36$ and $\xi_c = 0.84$, of course still consistent with the same ξ. In both cases ξ_G and ξ_c have reasonable values, as in general a proportionality constant between two correlated quantities is expected to be of the order of unity; if not, then some physical reason hidden in the concerned correlation should account for its actual order of magnitude. Actually the factor ten just introduced is not accidental, although it appears at the moment arbitrary and unjustified; its physical meaning will be highlighted in the next section. So are of interest the following values

$$M_u = 10m_u; \quad \xi_G = 0.36; \quad \xi_c = 0.84; \quad \mathrm{v}_u = 0.3c. \qquad (2,6)$$

These estimates imply that v_x of (2,4) takes the meaning of recession velocity v_u of today's universe boundary, being specifically calculated via Δr_u at our current time Δt_u. Yet there is no reason to think that the ratio $\Delta r/\Delta t$ is necessarily constant; so (2,4) prospects in general a variable expansion rate controlled by this ratio at different ages of the universe. Moreover, since v_u should reasonably depend also on the amount of mass within the universe, one expects a link between Δr_u and m_u or more likely M_u; in effect this conclusion will be confirmed in the next section.

At this point, therefore, the first target of the present model is to highlight how $\mathrm{v_u}$ is related to M_u via Δr_u, see in particular the next equation (3,3) that is the key together with (2,5) to link Δr_u and Δt_u to M_u. The model is described implementing first these today data, useful to assess the results, then it is also extended to past times when necessary. For reasons that will be clear soon, it is useful to begin with the matter era. The starting points of the present paper are not the general relativity and the Friedmann equations, but the quantum equations (2,1). The paper aims to check the effectiveness of this approach to formulate a possible model of universe. The worth of the present approach relies in particular on the fact that just (2,1) have been proven suitable to link the roots of the quantum mechanics to that of the special and general relativity [13,14].

3 Physical background of a possible model of the universe

According to (2,1) and positions (2,2), the key quantities of the present paper are not r_u and t_u, but the ranges $\Delta r_u = r_u - r_0$ and $\Delta t_u = t_u - t_0$. Let r_u be the current coordinate of the boundary of the universe at the time t_u, respectively defined with respect to an arbitrary initial value r_0 at the arbitrary time t_0. As previously emphasized, these latter coordinates are in turn fixed in an arbitrary space-time reference system R. Once accepting the quantum approach shortly introduced in section 2 to describe the universe as a quantum system, however, both r_0 and t_0 are deemed unknown and unnecessary to infer the eigenvalues of the physical observables, described instead by Δr_u and Δt_u only; moreover no particular R is specifiable, in agreement with one of the basic hypotheses of the relativity according which all reference systems are equivalent to describe the physical systems. If the uncertainty ranges only have physical meaning to define the quantum eigenvalues describing the observables, as shortly sketched in section 2, then this kind of universe has no defined center; this latter should be determined with respect to the origin of R, which however is undefined and indefinable itself like r_0 and t_0. Hence the physical universe is a space-time shell between the radii r_0 and r_u that define Δr_u. As the same holds for the time, the beginning of time defining the cosmological space-time is conceptually unidentifiable; it could be $t = 0$ or $t = t_0$ or any intermediate time. Strictly speaking, Δr_u and Δt_u only characterize the actual physical features of today's quantum universe. It means that r_0 and t_0, and in an analogous way ε_0 and p_0 of the respective ranges, characterize a pre-universe only; i.e. they are precursors of the space-time quantum ranges of (2,1) to which are actually related the physical observables of the universe. In fact, the following considerations will confirm the idea that trying to determine the initial values r_0 and t_0 is in fact inessential. The starting point of the present model is introduced as follows. Consider $\Delta p_r = n\hbar/\Delta r$ putting $\Delta p_r = h/\lambda_r - p_0$: coherently with Δr, also Δp_r defines an allowed range of local radial momenta falling between h/λ_r and p_0, both arbitrary. This equation yields in particular, specifying $\Delta r = \Delta r_u$,

$$n\lambda_u = 2\pi\Delta r_u; \quad \lambda_u = \lambda_r\lambda_0/(\lambda_0 - \lambda_r); \quad \lambda_0 = h/p_0. \qquad (3,1)$$

Whatever λ_0 might be, λ_r introduces a new wavelength λ_u; this result has in principle general valence because of the fundamental character of (2,1). For instance (3,1) imply a condition well known in quantum mechanics: an integer number n of wavelengths λ_u around a circumference corresponds to steady electron waves around a nucleus, in agreement with the quantization here introduced just by n. As λ_u has been defined without specifying the nature of the wave it characterizes, let us concern the particular case of a steady electromagnetic wave of wavelength λ_u traveling on the surface of a sphere. The assumption $r_0 \ll r_u$ brings thus to mind a hyperspherical four dimensional closed universe of radius Δr_u surrounded by a light wave running around any diametric circumference. This preliminary standpoint suggests in turn a possible hypothesis about its hypervolume and hypersurface

$$V_u = (4\pi/3)\,\Delta r_u^3; \qquad A_u = 4\pi\Delta r_u^2 \qquad (3,2)$$

filled with an amount of matter such to fulfill both (2,3) and (3,1). This also suggests regarding the universe consistent with the condition of "maximum growth efficiency", i.e. like a supermassive black hole; in effect, the previous considerations show that this conclusion is compatible with the analysis of the physical dimensions of G. Usually a black hole is allowed to form when any system, e.g. a star of sufficient mass at the end of its life cycle, collapses down to a critical radius fulfilling (2,3); so is seemingly surprising an expanding universe regarded as a supermassive black hole. Yet there is no physical reason to think that in general the shrinking process is the distinctive condition allowing a black hole; this usual idea implemented to explain observable events occurring inside the universe cannot be extrapolated to the behavior of the whole universe itself. Indeed Δr_s has been inferred via the physical definition of G simply exploiting (2,1), regardless of any specific reference to collapse events. Actually the present hypothesis seems reasonable for a growing universe, whose main requirement is to prevent mass and radiation energy losses outside it that could avert its possible evolution. According to the Hawking mechanism based on the vacuum polarization in the presence of a strong gravity field, a black hole inside the universe is able to split a couple of virtual particles generated by vacuum quantum fluctuation; it captures one of them, while releasing the other that thus appears as an ordinary particle. Outside the universe however this mechanism does not hold, as the concept of vacuum is replaced by that of "nothing". So no energy can escape outside Δr_u. The universe is thus a closed box unobservable from an external observer possibly existing. This point of view is assessed preliminarily by introducing the Schwarzschild range (2,3) and identifying $\Delta r_s \equiv \Delta r_u$ and $m \equiv m_u$; this position yields $\Delta r_s = 4.5 \times 10^{25}$ m, which is not very far from the estimated literature radius of the universe. Considering however that m_u quoted above is surely underestimated, as already emphasized, it is not surprising a value of Δr_s smaller than the expected Δr_u consistent with (2,3). Trust thus to the size of Δr_u and try to replace m_u with a value $M_u > m_u$ defined by

$$\Delta r_u = 2M_uG/c^2; \qquad (3,3)$$

one finds

$$M_u = 3 \times 10^{53}\ \text{kg}; \qquad M_u = m_u + m_? \approx 10m_u \qquad (3,4)$$

i.e. a total mass higher than the literature estimate of the visible m_u, as anticipated in section 2. This equation includes both the visible mass m_u plus a further contribution $m_?$ to be

explained next. Actually nothing excludes in principle the hypothesis (3,3), which in fact can be checked in several ways. So in the following M_u only, and not m_u, will be implemented. Estimate with the help of (3,2) and (3,3) the average density of the universe

$$\rho_u = \frac{3c^6}{32\pi M_u^2 G^3} = \frac{3}{8\pi G}\left(\frac{c}{\Delta r_u}\right)^2 \qquad (3,5)$$

which justifies why this paper starts just from the so called matter controlled era.

The most direct consequence of (3,3) is the Hawking entropy. Define first the circular frequencies of a light wave trapped by gravity around the border of the universe as

$$\omega_n = n\omega_u; \qquad \omega_u = c/(2\pi\Delta r_u)$$

in agreement with (3,1); so the boundary layer of the universe is marked out by the allowed frequencies of the electromagnetic field surrounding the total mass M_u, whose energy ε_ω is given by

$$\begin{aligned} \omega_n &= 1.1n \times 10^{-19}\ \mathrm{s}^{-1} \\ \varepsilon_\omega &= \frac{n\hbar c}{2\pi\Delta r_u} = 1.2n \times 10^{-53}\ \mathrm{J}. \end{aligned} \qquad (3,6)$$

Then let us concern also the total energy $\varepsilon_u = M_u c^2$ due to the whole amount of mass present in the universe. Since one expects that bulk energy ε_u and surface energy ε_ω should be someway correlated, the simplest hypothesis is to introduce a dimensionless proportionality factor σ_{H} such that $\varepsilon_u = \sigma_{\mathrm{H}}\varepsilon_\omega$. To infer the physical meaning of σ_{H}, calculate the mean values of this equation, which reads $<\varepsilon_u> = \sigma_{\mathrm{H}} <\varepsilon_\omega>$. Clearly $<\varepsilon_u> \equiv \varepsilon_u$. The standard way to calculate $<n\hbar\omega_u>$ via the partition function is well known; noting that $\hbar\omega_u \ll k_B T$ is verified for T down to values of the order of 10^{-28} K, one finds $<n\hbar\omega_u> \approx k_B T$. So $k_B\sigma_{\mathrm{H}}$ defined by an energy over a temperature can be nothing else but entropy. With the help of the Plank length $l_P = \sqrt{\hbar G/c^3}$, one finds indeed thanks to (3,2) and (3,3)

$$\sigma_{\mathrm{H}} = \frac{<\varepsilon_u>}{<n\hbar\omega_u>} = \frac{A_u}{4l_P^2}; \quad \hbar\omega_u = \frac{\hbar c}{2\pi\Delta r_u}; \quad \varepsilon_u = \frac{c^4}{G}\frac{\Delta r_u}{2}.$$

In effect, σ_{H} coincides just with the well known Hawking surface entropy in Boltzmann's units.

Before discussing further evidences to support the idea of black hole-like universe, as concerns in particular the value of M_u hypothesized here, let us implement the right hand side of (2,1): one finds $\Delta\varepsilon_u = \hbar/\Delta t_u$, whose physical meaning is clearly that of energy uncertainty range within which is defined the energy ε_u of the universe. Moreover, multiplying both sides by M_u, one finds

$$\Delta\varepsilon_u = \frac{\hbar}{\Delta t_u} = 2.4 \times 10^{-52}\ \mathrm{J}; \quad \Delta p_u = \sqrt{M_u \Delta\varepsilon_u} = 9\ \mathrm{kg\ m/s}.$$

So the uncertainty range of the momentum p_u of the universe has size of the order of the Planck momentum. The fact that the size of $\Delta\varepsilon_u$ is very narrow means of course that ε_u, whatever its value might be, is defined almost exactly. It is interesting to implement this result via the definition of G. Replace m with M_u and $\Delta t_u = \hbar/\Delta\varepsilon_u$ in the second (2,5); one finds thus $\Delta\varepsilon_u = \hbar c^3/(\xi_c^3 G M_u) = 1.4\xi_c^{-3} \times 10^{-52}$ J. Therefore $\Delta\varepsilon_u$ here calculated with $\xi_c = 0.84$, i.e. with the same value of (2,6), agrees with that obtained here directly from (2,1) via the age of the universe only. So this result on the one hand supports the value of M_u previously found, on the other hand it also confirms that the physical dimensions of G actually summarize the quantum features of the universe.

Owing to (3,3), the second (2,4) reads

$$\mathrm{v}_u = c^2\frac{\Delta t_u}{\Delta r_u} = \frac{c^4}{G}\frac{\Delta t_u}{2M_u} \qquad (3,7)$$

whose numerical value coincides of course with that of (2,6). According to (2,5), an increasing ratio $\Delta t_u/M_u$ means a smaller mass at Δt_u and thus a greater $\mathrm{v_u}$, as it is natural to expect.

To implement further these considerations, note that $\sqrt{\rho G}$ yields a frequency; so, replacing ρ with ρ_u of (3,5), one finds

$$\sqrt{\rho_u G} = 2.4 \times 10^{-19}\ \mathrm{s}^{-1}. \qquad (3,8)$$

This value is nicely twice the ground value of (3,6), even though calculated via G only and regardless of the condition (3,1); i.e. it requires $n = 2$. This result has a remarkable physical meaning that will be highlighted later. After having examined the physical meaning of the ratio $\hbar\omega_n/\varepsilon_u$ let us consider now the ratio $\hbar\omega_n/\Delta\varepsilon_u$: we emphasize that the deviation of M_u from the visible mass m_u is controlled by the constrain between (3,6) and (2,4), i.e. between the surface energy $\hbar\omega_{n=2} = \hbar c/(\pi\Delta r_u)$ of the electromagnetic wave surrounding the universe and the uncertainty energy range $\Delta\varepsilon_u = \hbar c^3/(\xi_c^3 G M_u) = \hbar/\Delta t_u$ of the bulk universe; indeed with the help of (3,3) and (3,6) we obtain

$$\omega_n \Delta t_u = \frac{nc\Delta t_u}{2\pi\Delta r_u} = \frac{n\mathrm{v}_u}{2\pi c}; \qquad \frac{\hbar\omega_n}{\hbar/\Delta t_u} = \frac{n}{2}\frac{\xi_c^3}{2\pi} \approx 0.05n \qquad (3,9)$$

according to the values (2,6), which yields $\hbar\omega_{n=2}/(\hbar/\Delta t_u) \approx 0.1 = m_u/M_u$. This result is crucial to understand the physical meaning of $m_?$, as highlighted in section 4.

Consider now that the ratio $c/\Delta r_u$ of (3,5) has physical dimensions time^{-1}; thus it is definable in general as $\dot{a}/a$, being a a function of coordinate and time. It is known that Δr_u^{-1} describes the local curvature of a surface; so $c/\Delta r_u$ must be actually expressed as $(\dot{a}+b)/a$ via an additive constant b, without which the curvature of the universe would tend to zero merely for a tending to a constant. Instead it seems more sensible to think that even for constant Δr_u the curvature becomes con-

stant itself, but not necessarily equal to zero. So (3,5) reads

$$\begin{aligned}\frac{8\pi\rho_u G}{3} &= \left(\frac{\dot a}{a}\right)^2 + \left(\frac{b}{a}\right)^2 + \frac{2\dot a b}{a^2}\\ \frac{c}{\Delta r_u} &= \frac{\dot a}{a}\left(1+\frac{b}{\dot a}\right)\end{aligned} \qquad (3,10)$$

i.e., more expressively,

$$\frac{8\pi\rho_u G}{3H^2} = \frac{\rho_u}{\rho_c} = 1 + \frac{b}{\dot a}\left(2+\frac{b}{\dot a}\right);\quad \rho_c = \frac{3H^2}{8\pi G};\quad H = \frac{\dot a}{a}.$$

Despite the quantum approach has been carried out regardless of the general relativity, the conclusion is that $b/\dot a < 0$ or $b/\dot a = 0$ or $b/\dot a > 0$ depending on the ratio ρ_u/ρ_c; either sign of $b/\dot a$ depends on that of $\dot a$ and b controlling the curvature according to (3,10). Calling $b = \pm c$ and $\Lambda = \mp 6H/(ac)$ the right hand side reads $H^2 + (c/a)^2 - \Lambda c^2/3$, i.e. this equation reduces to the Friedmann equation; H is the Hubble parameter and Λ the cosmological constant. The implications of the Friedmann equation, as concerns in particular the parameter k, are so well known that a detailed discussion of (3,10) is superfluous. We emphasize the crucial role of (3,3) to obtain directly from (3,5) this result, which however compels automatically accepting here $\rho_u/\rho_c > 1$ once having hypothesized since the beginning a closed universe with hyperspherical geometry. If this inequality is such that $\rho_u/\rho_c \gtrsim 1$, then the previous considerations are consistent with an almost Euclidean closed universe, in which case

$$\frac{b}{\dot a}\left(2+\frac{b}{\dot a}\right) \gtrsim 0. \qquad (3,11)$$

This is verified by $0 < b \ll \dot a$ and $b/\dot a \gtrsim -2$. Now, after having preliminarily verified the hypothesis (3,3) suggested by (3,1), let us check also the self-consistency of the considerations hitherto exposed examining once more $c/\Delta r_u$.

It is reasonable to think Δr_u proportional to the age Δt_u of the universe; so it is possible to write a series expansion defining Δr_u as $\Delta r_u = \sum_{j=1} a_j (cf)^j$, where $f = f(\Delta t)$ is an appropriate function of time to be defined and a_j are constant coefficients of the series. Rewriting more conveniently this series as $\Delta r_u = a_1 cf\varphi$, where $\varphi = 1 + a_2 cf/a_1 + a_3(cf)^2/a_1 + \cdots$, one expects that a_1 of the first order term should be close to the unity for the aforesaid reasons. Implement once again the physical dimensions of G similarly as done before and put in particular $f(\Delta t) \equiv \Delta t_u$; if this position is correct, then $\Delta r_u = a_1 c\varphi\Delta t_u$ with $\varphi \approx 1$ yields $a_1 \approx 2c/\xi_c^3$. On the other hand ξ_c of (2,6) has been calculated in order to fit the numerical value of $G = c^3\Delta t_u/(\xi_c^3 M_u)$ of (2,5), which results also in agreement with that of (3,9); as this equation of G reads $\Delta r_u = (2c/\xi_c^3)\Delta t_u$ with the help of (3,3), one finds at the first order $a_1 \approx 2c/\xi_c^3$ and thus $\Delta r_u \approx (2c/\xi_c^3)\varphi\Delta t_u$. Also this result agrees with the previous estimate of ξ_c defining $\Delta r_u/\Delta t_u$: in effect from (2,4) and (2,6), $\Delta r_u = (c/\xi)\Delta t_u$ compares well with $\Delta r_u = (2c/\xi_c^3)\Delta t_u$ because the values (2,6) verify $\xi^{-1} = 2/\xi_c^{-3}$. This confirms that effectively $\varphi \approx 1$. Hence defining

$$H_0 = \frac{1}{\varphi\Delta t_u}$$

one finds with $a_1 \approx 2c/\xi_c^3$ and once more the given value of ξ_c

$$H_0 = \frac{2c}{\xi_c^3\Delta r_u} = 2.4\times 10^{-18}\ \mathrm{s}^{-1}.$$

So at the first order H_0 coincides with Δt_u^{-1}; moreover the second (3,10) yields $H(1+b/\dot a) = \xi/(\varphi\Delta t_u)$, i.e. $1+b/\dot a \approx \xi H_0/H$ and thus $1 + b/\dot a \approx 1$ in agreement with (3,11). The present estimate of H_0 fits well the average value of the Hubble constant, which according to recent measurements falls in the range $(2.2 \div 2.6)\times 10^{-18}\ \mathrm{s}^{-1}$.

These results justify the advantage of introducing the present quantum model with the matter era; once having estimated H_0 and inferred the Friedmann equation, it is easy to describe also the radiation controlled era as shown below.

It is worth emphasizing the strategy of the present approach. The standard way to infer cosmological information is to find the solution of the gravity field equations and next to implement the Friedmann solutions: these equations provide information about the open or closed geometry of the universe. Here a different approach has been followed. The quantum equations (2,1) have been implemented since the beginning to introduce the wavelength λ_u and formulate by consequence the concurrent hypothesis (3,3) about a possible geometry of closed universe; thereafter this preliminary idea has been checked to infer (i) the Hawking entropy, (ii) the link between mass density and curvature radius of the universe, (iii) to obtain a Friedmann-like equation and (iv) to estimate the Hubble constant. Moreover, exploiting the same approach outlined in section 2 for the angular momentum, the factors that define the physical dimensions of G allowed to correlate correctly size, age and mass of the universe. The remainders of this paper aim to implement these preliminary ideas to show that further reasonable results are inferred hereafter.

3.1 The matter era

Let us estimate the average mass and energy densities ρ_u and $\eta_{in} = \rho_u c^2$ of the universe, which result to be with the help of (3,2) and (3,5) of the order of

$$\begin{aligned}\rho_u &= \frac{M_u}{V_u} = \frac{3c^2}{2A_u G} = 8.7\times 10^{-28}\ \mathrm{kg/m^3}\\ \eta_{in} &= \frac{M_u c^2}{V_u} = \frac{3c^4}{2A_u G} = 7.8\times 10^{-11}\ \mathrm{J/m^3}.\end{aligned} \qquad (3,12)$$

These values reasonably agree with that calculated in a very different way in [18]; the corresponding "non-visible" energy density is instead of the order of

$$\eta_? = 3m_? c^2/(4\pi\Delta r_u^3) = 7\times 10^{-11}\ \mathrm{J/m^3};\qquad m_? \approx 9m_u.$$

The ordinary visible mass of the universe is about 10% of the total mass only, whereas the remainder mass $m_?$ accounts for the 90% gravitational effect responsible of the black hole-like behavior of the whole universe. The average density ρ_u hides the physical nature of the actual total mass. Moreover, besides m_u of visible stars, M_u consists of a preponderant contribution $m_?$ of different physical nature: for instance all black holes possibly existing in the universe, or interstellar gas and dust, or free elementary particles, and so on including also the so called dark mass. A complex system of particles contributes to M_u, whose actual nature is however not explicitly concerned in neither of (3,12). According to some theories the elements were formed inside the stars by neutron bombardment of light nuclei and subsequent β decay, e.g. [19], other authors believed instead that elements were formed during the early stages after the big bang, e.g. [20]; more recently other authors returned to their formation inside the star by virtue of several nuclear processes [1]. Despite (3,12) waive specific information about the actual composition of M_u, the assumption of large scale homogeneity and isotropy of the universe supports the effective physical meaning of average ρ_u. Moreover the concept of quantum delocalization introduced by (2,1) stimulates itself the idea of average mass spreading uniformly throughout the universe likewise as the energy field of light radiation. This idea is useful to link the matter era to the earlier radiation era. It will be emphasized in the next subsection 3.2 that the radiation field, almost mono-chromatic at the beginning of the radiation era, turned into a more complex spectrum of wavelengths because of the concurrent expansion of the universe; so quantum fluctuations and possible events of constructive interference, statistically allowed to occur anywhere in the radiation field, promoted favorable conditions to form local couples of virtual particles uniformly distributed in the available volume of the early universe. It is known indeed that proton and antiproton virtual couples are formed by vacuum fluctuations and high order two-photon interactions during photon fluctuations able to generate fermion-antifermion pairs [21]. So it seems reasonable to guess that this mechanism triggered the evolution of the early radiation field to couples of virtual particles continuously annihilating and re-materializing up to the later formation of colder real matter. Some considerations on this point will be shortly sketched in the appendix. For the purposes of the present paper, however, it is enough to acknowledge that today's ρ_u corresponds on average to about one half proton mass per cubic meter of universe and that (3,12) hold identically while considering the mass of antiprotons. Despite this idea is mere statistical abstraction, (3,12) are useful for the purposes of the present model; they implement the assumed homogeneity and isotropy of the universe in its strongest form possible. Even with such information only, i.e. whatever the actual abundances of the j-th elements of mass m_j might be today within each unit volume of universe, it is possible to introduce: (i) an elementary volume V_0 physically located anywhere and defined as that containing on average one proton or one antiproton and (ii) a linear combination $m_p = \sum_j a_{ij} m_j$ that accounts via the local coefficients $a_{ij} = a_{ij}(x_i, y_i, z_i, t, m_j)$ for the actual composition of real matter progressively formed everywhere after the radiation era. These coefficients weight the time profile of the local effective abundances, e.g.: they are null if the pertinent coordinates of a_{ij} correspond to an empty volume of universe where $m_j = 0$, moreover all a_{ij} were equal to zero during the early radiation era, and so on. Since the local coordinates are conceptually disregarded by (2,1) and positions (2,2), however, let the indexes i and j number respectively the N_{in} elementary volumes V_{0i} of the universe and the various elements therein formed a time range Δt after its birth. The abundances are subjected to the boundary condition of the first (3,12); for instance, at today's Δt_u this point of view is summarized by the sums

$$\begin{aligned} \rho_u &= \frac{1}{V_u} \sum_{i,j} a_{ij} m_j \\ a_{ij} &= a_{ij}(V_{0i}, \Delta t, m_j) \\ \sum_{i,j} a_{ij} m_j &= N_{in} m_p . \end{aligned} \qquad (3,13)$$

The first two equations emphasize the local composition of ρ_u, the last one fits in particular the condition of today's average density. In fact (3,13) regard the universe as a lattice, whose elementary cells are the volumes V_{0i} uniformly occupied by one proton or one antiproton of every virtual couple with equal probability. Each cell is therefore a possible allowed state for either of them, i.e. the universe is statistically described by a total number $N_{in} = V_u/V_0 \approx 1.7 \times 10^{80}$ of degenerate states corresponding to η_{in}; also, since by definition each V_0 contains on average one proton mass, $m_p N_{in} = M_u$. So according to (3,12) the energy levels ε_{V_0} of one proton or one antiproton in the respective V_0 states are $m_p c^2/2$ and $m_{\bar{p}} c^2/2$, i.e.

$$\varepsilon_{V_0} = 7.8 \times 10^{-11} \text{ J}; \qquad V_0 \approx 2 \text{ m}^3 \qquad (3,14)$$

in order that effectively $M_u/V_u = m_p/V_0$, in agreement with (3,12). Of course ε_{V_0} includes also the interaction energy between charges in different cells, e.g. that of couples of all virtual particles possibly generated together with energetic protons and antiprotons; this is possible because $M_u c^2$ involves the visible mass energy $m_u c^2$ plus the contribution of $m_? c^2$. Note eventually that despite $M_u c^2$ results statistically equivalent to the sums

$$\sum_i (\varepsilon_{V_0 prot} + \varepsilon_{V_0 antiprot}) = N_{in} m_p c^2/2 + N_{in} m_{\bar{p}} c^2/2 \qquad (3,15)$$

over all the elementary volumes V_0, it will be shown later that an effective entropy driven mechanism in fact marked the transition from the radiation era to the matter era; so the sum

of (3,15) reads actually

$$\sum_i (\varepsilon_{V_0\,prot} + \varepsilon_{V_0\,antiprot}) = N_{in} m_p c^2. \qquad (3,16)$$

Before describing this mechanism, the results so far obtained are summarized as follows: (i) each cell is in fact an allowed state for one proton or one antiproton; (ii) (3,14) represents the excitation energy necessary to remove either of them from its own V_0 and leave behind an empty cell; (iii) the latter represents a vacuum state, whereas either particle present in V_0 defines an occupied state.

To highlight the physical meaning of these points, consider an arbitrary mass m at the boundary of the universe. The shell theorem shows that the gravity force acting on m is that due to M_u regarded in the ideal center of a spherical body; so is accordingly calculated for a radius Δr_u its energy $\varepsilon = GM_u m/\Delta r_u$ that, exploiting once again (3,3), reads also $\varepsilon = mc^2/2$. If for instance m represents the mass of one proton or one electron, $m_p = 1.7 \times 10^{-27}$ kg and $m_e = 9.1 \times 10^{-31}$ kg, then one finds

$$\begin{aligned} \varepsilon_p &= G\frac{M_u m_p}{\Delta r_u} = \frac{m_p c^2}{2} = 1.0 \times 10^{-10}\ \text{J} \\ \varepsilon_e &= G\frac{M_u m_e}{\Delta r_u} = \frac{m_e c^2}{2} = 5.4 \times 10^{-14}\ \text{J}. \end{aligned} \qquad (3,17)$$

The second (3,17) emphasizes that if the volume V_0 would be occupied by one electron with its own energy level $m_e c^2/2$, then V_0 would represent a possible state for this electron. To clarify where anyway does m come from, note that at today's Δt_u the proton energy level ε_{V_0} inside any state V_0 of the bulk universe, (3,14), is equal to the energy ε_p, (3,17), of one proton at the boundary of the universe. So

$$\varepsilon_p = \varepsilon_{V_0}. \qquad (3,18)$$

This equation in fact reads $c^2/2 = M_u G/\Delta r_u$, which is nothing else but (3,3). Thus (3,18) and the first (3,17) do not depend on the proton mass, and hold whatever else m_p might represent. Moreover neither the analytical form of ρ_u nor that of η_{in} introduce explicitly m_p. Rather, the latter introduces the mere Planck force c^4/G acting on the total surface A_u of the universe. There are two reasons why the average values defined by (3,12) and (3,13) have importance for the following discussion: on the one hand, the right side of (3,12) links correctly energy density and pressure; on the other hand, being known that the pressure of a perfect gas is 2/3 of its energy density, the second (3,12) suggests regarding η_{in} in each volume V_0 as due to a proton/antiproton gas occupying uniformly all bulk states of the universe. As this average pressure appears to be a physical property of all elementary volumes V_0, then the internal pressure that characterizes the whole universe results to be, again via (3,3),

$$P_{in} = \frac{2}{3}\frac{M_u c^2}{V_u} = \frac{c^4}{A_u G} = 5.6 \times 10^{-11}\ \text{Pa}. \qquad (3,19)$$

The fact that even P_{in} does not depend explicitly on m_p suggests that (3,12) have actual physical meaning. The factor 2/3, numerically irrelevant in the frame of the order of magnitude estimates proposed here, is however conceptually significant to check the physical meaning of (3,12). Taking into account (3,16), (3,19) reads

$$P_{in} V_u = \frac{2}{3} E; \qquad M_u c^2 = E = N_{in} m_p c^2.$$

The surprising fact is that the mere definition of energy density, without any additional hypothesis, portrays the whole universe as a container full of quantum or classical gas, whose mass M_u exerts Planck force against its inner boundary; indeed the first equation holds for Boltzmann, Bose and Fermi statistics, which confirms that effectively any kind of quantum or classical particle, thus why not the proton, is compatible with M_u without affecting the validity of (3,19). Furthermore this picture holds at any time, because the surface A_u can be replaced by any A likewise related to the pertinent M/V whatever the numerical value of the ratio might be. Formally this is justified by the second equation, where E resulting from $M_u c^2$ is also associated to a number N_{in} of proton masses fulfilling the global energy conservation. Yet the simple equivalence matter/energy does not seem enough to explain why chunks of matter like asteroids or stars or cosmic powder could mimic the pressure of a proton gas of equivalent total mass filling uniformly the universe. This is however a classical way to think the universe. More stimulating appears in this respect the quantum character of the present model. First of all, the couples proton/antiprotons have been guessed as mere numerical hint due to the average value of the mass resulting in (3,12); but in fact any gas could be consistent with (3,19), which indeed does not make explicit reference to m_p. The chance that any gas mixture could contribute to E is a step towards introducing the actual existence of chemical abundances symbolized by various m_j; the first (3,13) merely means that the degenerate proton or antiproton energy levels $m_p c^2/2$ split into a complex system of non-degenerate energy levels describing the local bound states of cosmic matter. From this point of view, the energy conservation between two different systems of quantum energy levels appears more pertinent: since in principle one level could split into several non-degenerate levels in an infinite number of ways, the energy conservation appears as essential boundary condition to calculate the latter from the former, rather than a mere statistical abstraction. More significant is however the dual wave/corpuscle behaviour of matter. A body of real matter is superposition of waves to form a group in principle spreading from minus infinity to infinity but with a maximum probability of being somewhere: the amplitude of the wave packet rapidly decreases at the edge of a region that determines the most probable position and the finite extent of the body, whose possible motion is nothing else but the group velocity of the wave packet. It is known that the electromag-

netic waves exert a pressure, whence the photon gas physics: why not to think the same about delocalized matter waves, according to (2,1)? If so, then the matter era began when matter waves started to appear in the pre-existing field of electromagnetic radiation according to the mechanism [21]. The appendix gives some more hints on this topic.

On the one hand these considerations are interesting because P_{in} controls the expansion of the universe, as it will be shown below; on the other hand the idea of V_0 bulk states allowed to protons and antiprotons, although suggested by the numerical values of (3,12) only, is attracting because it links radiation era and matter era, at the beginning of which couples of matter/antimatter particles were in fact formed. Anyway the significant conclusion is that (3,17) to (3,19) skip m_p and thus can be further implemented in the following regardless of whether the volumes V_0 are really occupied by protons or any other mass.

Exploit (3,19) to infer the average temperature T related to P_{in} in V_0. Here $T \approx E/(N_{in}k_B) = m_pc^2/k_B$ helps to estimate the average temperature in each elementary volume V_0; one finds $T \approx 10^{13}$ K. This estimate fulfills the usual statistical meaning of temperature, as the proton here concerned has a statistical meaning itself. To better assess this result consider the pressure P of an ideal gas of molecular weight M_{mol} and average density ρ in the volume V_0, so that $\rho = PM_{mol}/RT$. Exploiting (3,12) and (3,19) at the time Δt_u to express $\rho \equiv \rho_u$ and $P \equiv P_{in}$, one finds $M_u/V_u = (2M_uc^2/3V_u)M_{mol}/RT$, i.e. $T = 2c^2M_{mol}/3R$. Hence T is explicitly related to the specific M_{mol} only, regardless of the time Δt_u and related universe volume V_u. A uniform distribution of hydrogen in each V_0, i.e. $M_{mol} = 10^{-3}$ kg, estimates again $T \approx 10^{13}$ K, in agreement with that inferred directly from m_pc^2/k_B. Even the formation of hydrogen will be justified in the subsection 3.4 as a consequence of the step from (3,15) to (3,16). This large value is enough for protons to form further couples of virtual photons and fermions/antifer-mions; this supports the idea that effectively the protons early formed trigger the successive energy balance in V_0 qualitatively indicated in (3,13).

The previous ways to estimate T refer to the time where early hadrons began to form everywhere in the radiation field of such universe and indicate a temperature corresponding to a uniform distribution of virtual couples occupying the available states at the end of the radiation era. The same equations could in principle estimate the local T even during the subsequent matter era, when the bombardment with energetic neutrons allowed forming heavy elements; yet the concurrent clustering of matter determined a structure of the universe locally inhomogeneous, so at that later time a unique average T does no longer make sense. Actually both time and volume of the universe determine the value of M_{mol}. In particular, the expansion of the universe is crucial to determine the time profile of T after the radiation era: the hypothesis (3,3) requires $M/\Delta r = const$, which also compels that $M/\Delta r^3$ is a decreasing function of time for increasing Δr. So an increasing fraction of empty zones of the universe corresponds in principle to a global decreasing value of T; the calculation of the respective temperatures is not as immediate and straightforward as in the previous case, characterized by a uniform distribution of a unique kind of early particles. In this case both local coefficients a_{ij} and atomic weights of the elements m_j must be known: the sums of (3,13) are related to the abundances within the various volumes V_{0i} of cosmic objects, characterized by the different kinds of elements and local coefficients a_{ij}, and to empty parts of the universe.

A question arises now: did (3,3) and (3,18) hold even in the past? In fact there is no reason to suspect that this condition is an exclusive feature of the today space-time coordinates Δr_u and Δt_u, which indeed have nothing special with respect to any past or future Δr and Δt. The only necessary hypothesis to answer affirmatively is that the current V_0 grows together with the size of the universe, which is possible if its sizes are comoving distances. Otherwise stated, let V_0' be the past value of V_0 at any $\Delta r < \Delta r_u$ and $\Delta t < \Delta t_u$; we require $m_pc^2/2 = Mm_pG/\Delta r$, being M the past total mass. This requirement emphasizes the previous remarks: the actual nature of proton mass m_p is irrelevant as concerns (3,18), which holds thus whatever m_p stands for, i.e. whatever the relative element abundance of (3,11) in V_0 might have been at Δt. On the one hand $c^2/2G = M_u/\Delta r_u$ requires $M_u/\Delta r_u = M/\Delta r$ and thus $M = c^2\Delta r/2G$, i.e. the black hole condition held also in the past. On the other hand one expects that V_0' scales with $\propto \Delta r^3$, in order that it be definable even for the smaller universe sizes of the early matter era; so $V_0' = (\Delta r/\Delta r_u)^3V_0$, i.e. V_0' was reasonably much smaller than today's V_0. In this way multiplying both sides by N_{in} one finds $N_{in}V_0' = (\Delta r/\Delta r_u)^3N_{in}V_0$; since by definition $N_{in}V_0 = V_u$, (3,2) yield $N_{in}V_0' = (4\pi/3)\Delta r^3$, i.e. in the early hypersphere volume defined by Δr the number of elementary volumes and thus of states allowed to the new born matter was the same as today's N_{in}. In summary

$$M = \frac{\Delta r}{\Delta r_u}\, M_u; \quad V_0' = \left(\frac{\Delta r}{\Delta r_u}\right)^3 V_0; \quad N_{in} = const. \qquad (3,20)$$

What is important for the following discussion is that under reasonable assumptions the condition (3,18) could hold also in the past and that N_{in} was since the beginning fingerprint of our universe. (3,20) help to guess the size of the universe at the beginning of the matter era. It is instructive to proceed stepwise calculating Δr and V_0' by trial and error, i.e. assessing these quantities as a function of sensible values of M. If M would be the mass of one couple proton/antiproton only, then $\Delta r \approx 4.9 \times 10^{-54}$ m, which would mean a volume $V_0' \approx 2.9 \times 10^{-240}$ m^3, unrealistically smaller than the expected order of magnitude of Planck volume. This value of V_0' suggests an early number of virtual couples much higher than this. More reasonable results are obtained putting $V_0' \approx 4.2 \times 10^{-105}$ m^3 to estimate via the second equation the

order of magnitude of Δr, which results $\Delta r = 5.5 \times 10^{-9}$ m; with this range the first equation yields $M = 3.8 \times 10^{18}$ kg corresponding to about 2.2×10^{45} protons, i.e. about 10^{45} virtual couples proton/antiproton at the beginning of the matter era. Note that $Mc^2 = 3.4 \times 10^{35}$ J corresponds to an average fluctuation energy $\varepsilon_{fl} = 3.4 \times 10^{-10}$ J, i.e. 2.1 GeV, per virtual couple of matter particles newly created: this is the fluctuation energy of the radiation field able to create matter. It is interesting the fact that with the given choice of V_0' this result fits well the energy of a couple of protons, despite it has been calculated implementing M_u and Δr_u via (3,20) only; this supports the interpretation of (3,12). Supposing that on average each couple of photons generates one virtual couple of matter/antimatter, the fluctuation extra energy of radiation field increases the early Planck frequency of each couple of photons by about $\delta\omega = 3.4 \times 10^{-10}/\hbar = 3 \times 10^{24}$ s^{-1} to produce matter. The obvious conclusion of this section is to admit that before the time of mass production there was an earlier massless era, i.e. the radiation era.

3.2 The radiation era

Consider the density ρ corresponding to M and Δr of (3,20) by replacing M with $h/(\lambda c)$; in this way the total mass of the universe is expressed via the momentum h/λ of an electromagnetic wave propagating with velocity c. For simplicity we have assumed that the refractive index of the medium where the wave propagates is 1, although in principle this is an approximation only; the aforesaid gamma-gamma physics [21] predicts photon fluctuations resulting in charged fermion -antifermion pairs, leptons or quarks, which couple with the photons themselves. In the presence of electron-positron and proton-antiproton couples of particles that typically also form as a consequence of this kind of interaction, a refraction index equal to 1 is certainly an approximation; yet this is acceptable for the following reasoning and order of magnitude estimates. So the late $\rho_u = 3M_u/(4\pi\Delta r_u^3)$ of matter era reads $\rho^r = 3h/(4\pi\lambda c\Delta r^3)$ at the time Δt. A boundary condition for λ comes from the fact that the early electromagnetic radiation waves bounced between diametric distances $2\Delta r$ inside a sphere, i.e. still $\lambda = 2\Delta r/n$ with n integer according to eq (3,1); in this way steady waves were allowed to fill the universe at any time Δt. The internal bouncing of radiation is justified even admitting that the early stages of growth were allowed in non-equilibrium condition, owing to the rapid growth of the universe size, and without radiation energy loss unfavorable for the subsequent growth and evolution of the new-born universe. So λ was a function of time like Δr, i.e. the number n of allowed frequencies increased along with Δr; it seems reasonable to guess that an initial field almost monochromatic evolved towards a complex spectrum of steady wavelengths. Anyway the density of the universe in the radiation era reads

$$\rho^r = \frac{3nh}{8\pi c\Delta r^4} = \frac{3n\hbar}{4c\Delta r^4}.$$

while (3,3) reads $\Delta r = 2hG/(\lambda c^3)$; so the condition $\lambda = 2\Delta r/n$ yields $\Delta r = \sqrt{nhG/c^3}$. Hence increasing n means increasing Δr and the number of states allowed for the radiation field. So radiation density, radiation energy density and pressure during the radiation era read

$$\rho^r = \frac{3c^5}{4n\hbar G^2}; \qquad \eta^r_{in} = \frac{3c^7}{4n\hbar G^2}; \qquad P^r_{in} = \frac{c^7}{4n\hbar G^2}.$$

At the beginning of the radiation era, therefore, $\Delta r = \sqrt{hG/c^3}$ with $\lambda = \Delta r$ and $n = 1$ has the expected order of the Planck length with which in effect has been calculated the Planck volume V_0'. Moreover estimating hc/λ with λ of the order of the Planck length, $\approx 10^{-35}$ m, yields a temperature $T \approx hc/k_B\lambda$ of the order of 10^{33} K. The fact that this characteristic temperature is much higher than that estimated for the proton in today's V_0, confirms that actually the radiation era precedes the matter era. Putting Δr of the order of the Planck length, with $n = 1$ one finds $\rho^r \approx 4 \times 10^{96}$ kg/m^3 and $P^r_{in} \approx 10^{113}$ Pa and $\eta^r_{in} = 3.5 \times 10^{113}$ J/m^3; at this stage of evolution of the universe the energy $\varepsilon^r_{in} = (4\pi/3)\Delta r^3 P^r_{in}$ results about $\varepsilon^r_{in} \approx 1.7 \times 10^9$ J, to which corresponds a temperature of the order of $\varepsilon_{in}/k_B \approx 10^{32}$ K in agreement with that already estimated. Estimating an energy $k_B T \approx 1.3 \times 10^9$ J of the radiation field corresponding to this temperature, one finds $\omega^r = 1.6 \times 10^9/\hbar \approx 1.6 \times 10^{43}$ s^{-1} i.e. a radiation field with Planck frequency. These values correspond well therefore to the Planck pressure, energy, frequency and temperature.

So, trying to understand the physical meaning of these results beyond the numerical estimates, the radiation era was just after the very early time step of the creation of radiation just concerned; this initial step can be therefore nothing else but the Planck era. The huge internal pressure accounts for the rapid volume of the universe. Note that the value of ε^r_{in} is large, but not spectacularly high like P^r_{in} and η^r_{in}; these latter are due to the extremely small values of Planck volume. These ideas explain thus the subsequent beginning of the matter era, during which however the expansion mechanism of the universe was somehow different.

3.3 The universe expansion in the matter era

Comparing (3,17) and (3,14), it has been already noted the similarity between the gravitational energy ε_p of one proton at the boundary distance Δr_u and the energy ε_{V_0} existing within each V_0 just because of the presence of the proton itself. (3,20) have been accordingly inferred. If the proton, or whatever else its mass might actually represent, would be ideally removed from any volume V_0 internal to the universe and displaced to the boundary of the universe, the energy lost by V_0 is balanced by that transferred to the boundary; within the limits of the present order of magnitude estimates, there is no net gain or loss of energy in this ideal process. This suggests that creating a vacancy in the universe after ideally moving its average amount of matter per unit cell just to the external

boundary of the universe occurs at zero energy cost. Strictly speaking ε_p should have been calculated in principle writing $M_u - m_p$, the numerical difference being however completely irrelevant for one proton only. Actually this reasoning is extensible to describe a relevant number of protons regarded at the boundary; as $M_u/m_p \approx 10^{80}$, for a large number n_p of protons such that $1 \ll n_p \ll M_u/m_p$ still holds (3,18) because $M_u \approx M_u - n_p m_p$. This means that large numbers of protons are expected to contribute to this ideal transfer process, i.e. large numbers of empty cells are to be expected in the universe. Of course the comparison between ε_{V_0} and ε_p has statistical meaning only, despite the actual structure of the visible mass in the universe and even regardless of the local element abundances in the universe, hidden within the global value of M_u and still undisclosed when reasoning about the mere average distribution of M_u. The following remarks are useful at this point.

- There is no actual flow of protons moving inwards or outwards throughout the universe; the uncertainty in the most agnostic form of (2,1) requires any quantum particle completely delocalized everywhere in the whole universe. The diameter $2\Delta r_u$ is a quantum delocalization range inside which no information is conceptually allowed about the local position and dynamical variables of any kind of particle, proton or else. So any particle could be in V_0 or at the boundary simply provided that there are available allowed states; (3,18) merely compares the energies of protons in two different places where they could in fact be, i.e. everywhere because V_0 could be itself everywhere in the universe.
- Two states of equal energy are allowed to the proton: the bulk state in V_0 and the boundary state at the rim of the universe. A proton at the boundary state leaves behind an empty cell V_0, i.e. a hole in one of the bulk allowed states. In general occupied and empty states are possible in the bulk and at the boundary of the universe. The global electroneutrality is ensured by the identical chance statistically allowed to antiprotons too.
- Both ideal chances are possible in principle despite the black hole character of the universe: the protons do not escape far from the boundary, they remain "glued" on the boundary like any electromagnetic radiation possibly arriving up there from the bulk of the universe. The Hawking entropy supports this idea.
- The chance of either alternative is consequence of the second law of thermodynamics; these bulk and boundary chances concurrently possible for the protons increase their number of allowed states and thus their configuration entropy. This crucial point, which will be further concerned later, agrees with the fact that (3,17) describes identically the total mass M_u at the ideal center of the universe and the mass m_p at the boundary Δr_u apart or, vice versa, the mass m_p at the ideal center of the universe and the total mass M_u concentrated on a point at the boundary Δr_u apart; indeed, according to the considerations of section 2, the local position of any particle is physically meaningless because of the quantum delocalization within an uncertainty range. Either extremal configuration, in principle possible for the universe, is however unlikely by entropy considerations.
- If V_0 scales as described by (3,20), which is admissible as no restraining hypothesis has been made on it, then (3,18) previously introduced for the proton at the time Δt_u is unchanged at any $\Delta t < \Delta t_u$; moreover the number of states N_{in} is expected constant, as in effect it has been found.

These ideas encourage regarding the proton in V_0 as a sort of template that symbolizes the average behavior of real matter in any bulk state and at the boundary state; as previously remarked, this is certainly the strongest form to affirm the large scale isotropy and homogeneity of the universe. Actually particles and antiparticles with the same m_p concurrently formed after the radiation era have statistically the same probability of being found in the boundary state; if so, the initial configuration of coexisting protons and antiprotons uniformly occupying all available bulk states generates subsequently a boundary halo of virtual couples plus possible annihilation photons along with corresponding vacuum states and matter states in the bulk universe. This configuration change increases the total entropy of the universe. In particular, the surface entropy at the boundary of the universe consists of the Hawking term σ_{H} plus a contribution related to the configuration of boundary states shared with that of the bulk universe. The entropy will be considered in some more detail in the next section. It will be shown that the way of thinking based on the degenerate quantum states of the universe rather than on the multiplicity of states describing its actual structure of matter, helps formulating a possible growth mechanism of the universe. Usually growth and expansion are synonyms; the next section emphasizes why actually it is not so in the present model, where growth does not merely mean swelling.

3.4 The universe growth in the matter era

Let the bulk universe at an arbitrary time after the big bang consist of a number N_{out} of V_0 empty cells and a corresponding number $N_{in} - N_{out}$ of filled V_0 cells; the external boundary is thus a layer formed by N_{out} glued protons and antiprotons missing in the bulk. So even this statistical picture of universe is consistent with the existence of an empty part of the real universe and its real matter structure: correspondingly to the further redistribution of $N_{out}V_0$ and $(N_{in} - N_{out})V_0$ volumes, in principle located randomly in the total volume $N_{in}V_0$ available, clusters of matter tend to coalesce together by gravitational interaction: the vacuum corresponds indeed to the N_{out} residual holes left in between. Anyway, if clusters

of empty cells and clusters of occupied cells are numerous enough, then their random distribution within V_u is still consistent with the assumption of overall statistical homogeneity and isotropy. This seems indeed the case, as the number N_{in} of V_0 volumes has been estimated of the order of the Eddington number 10^{80}. The chance of introducing arbitrary numbers N_{out} and $N_{in} - N_{out}$ of cells brings the universe towards a situation of dynamical equilibrium between the former and the latter; yet this final configuration, somehow attained, could be imagined as the conclusion of a gradual process consisting of a first redistribution step $N_{in} - N'_{out}$ and N'_{out} of filled and empty cells, which in turn generates progressively a subsequent redistribution $N_{in} - N'_{out} - N''_{out}$ and $N'_{out} + N''_{out}$ of new filled and empty cells along with possible coalescence of cells still filled, and so on. This idea stimulates considering the dimensionless entropy of a current configuration, $\sigma_b = N_{in}!/(N_{out}!(N_{in} - N_{out})!)$, due to the fact that all transient configurations compatible with zero energy balance are equiprobable; the subscript b stands for "bulk". As σ_b has a maximum as function of N_{out}, the formation of bulk holes fulfills the second law until this maximum is reached. Let σ_b describe a transient configuration at a given time and $\sigma'_b = N_{in}!/(N'_{out}!(N_{in} - N'_{out})!)$ that at a later time; the latter is allowed if N_{out} and the subsequent N'_{out} fulfill $\sigma'_b > \sigma_b$. Hence, after an arbitrary numbers of steps, are formed as a function of time multiple clusters of matter aggregates subsequently attained and thus differently configured, together with a progressive modification of the empty space between them. At the dynamical equilibrium no net state exchange occurs. Of course σ_b and σ'_b neglect, for simplicity and brevity, the further contributions σ_{arr} and σ'_{arr} due to the ways to arrange the respective clusters of matter into actual universe structures; yet σ_b and σ'_b symbolize qualitatively the first conceptual step to understand the actual configuration of the universe. Clearly, by virtue of (3,13), the σ_{arr} driven final arrangements of filled cells are nothing else but stars or galaxies or flows of elementary particles or any other observable object. The existence of P_{in} related to the matter energy density agrees with and justifies the universe expansion, which however at this point still seems like a mere bubble blowing up by internal pressure effect. But just this point poses a further question: does the universe in the matter era expand freely or is it constrained by an external pressure P_{out} opposing to its expansion? In principle the expansion requires $P_{in} > P_{out}$, not necessarily $P_{out} = 0$: the force that pushes forwards the unit surface of universe boundary must simply overcome that possibly tending to pull it backwards, i.e. to squeeze the universe size towards a big crunch. If the former position is correct, then P_{out} tends to decrease the acceleration with which the universe expands. Yet, what does originate P_{out}? A possible answer relies just on the presence of protons and antiprotons at the boundary states of the universe previously introduced. The boundary here introduced is not mere spherical rim; in effect the plain idea of geometrical margin would be unphysical itself. More sensibly, the mobile contour of the universe is defined by a crowd of N_{out} virtual protons and antiprotons along with electromagnetic radiation trapped on a fading shell, recall the Hawking entropy. In fact the previous considerations propose in a natural way that the boundary should be a physical layer of finite volume and finite thickness; so the chance of defining an energy density η_{out} due to these particles seems the most straightforward way to define P_{out}. In this respect, the further chance of demonstrating that $P_{out} \neq P_{in}$ is important not only to infer information about the acceleration of the boundary of the universe, controlled by the net force $P_{in} - P_{out}$ per unit surface of boundary, but also to infer that the physical nature of the outer layer must be different from that inside the universe. Before assessing the importance of this conclusion as concerns the matter/antimatter ratio, let us examine two points: the expansion equation and the physical meaning of η_{out}, to which is related the pressure P_{out} equivalently as in (3,19). This external pressure could be likewise regarded as external force acting towards the center of the universe or resistance of the universe to increase the total surface of its boundary. The latter idea is more easily viable to introduce the existence of a boundary layer, whose thickness surrounds the universe and characterizes $\eta_{in} \neq \eta_{out}$; if the layer would have the same physical nature of the bulk vacuum, then the boundary should be at rest or steadily moving rather than accelerating. Let $\rho_u V_u c^2$ be the energy stored inside the universe; since today's universe expands, according to the first law its total energy E must also include a PV_u-like term. Let $\delta E = c^2\delta(\rho_u V_u) + P_{net}\delta V_u$ be the change δE of total energy during the time interval δt, where $P_{net} = P_{in} - P_{out}$ describes the net force pushing forwards the boundary. As no energy escapes outside of a black hole universe $\dot{E} = \dot{\rho} V_u c^2 + \rho \dot{V}_u c^2 + (P_{in} - P_{out})\dot{V}_u = 0$; so $\dot{\rho} + \rho \dot{V}_u / V_u + (P_{in} - P_{out})\dot{V}_u/(V_u c^2) = 0$. According to (3,20), the size of the elementary volume V_0 scales as Δr^3, i.e. like $V_u = N_{in} V_0$; then $\dot{V}_u / V_u = 3\dot{a}/a$, whence the well known result

$$\dot{\rho} + 3\,\frac{\dot{a}}{a}\left(\rho + \frac{P_{net}}{c^2}\right) = 0; \qquad P_{net} = P_{in} - P_{out} \qquad (3,21)$$

The notation emphasizes that the time derivative of the radius defines the change rate of a co-moving length. The excess of internal pressure means that the layer outside the boundary is slightly different from the bulk. Note that also a negative pressure P_{out} counteracting P_{in} has been introduced in this reasoning.

Regard the boundary as if it would be a material layer characterized by a contractive energy per unit surface $\gamma = \varepsilon_\gamma / l^2$ that opposes to its stretching during the expansion; for instance, this effect can be guessed thinking to the opposite charges of the particles/antiparticles that crowd the boundary surface. Anyway the total contractive energy of a spherical bubble having internal radius Δr_u and volume V_u is $\varepsilon_\gamma = 4\pi\Delta r_u^2 \gamma$. Moreover the Young-Laplace equation of such sur-

face tension-like model of boundary reads $P_{in}-P_{out} = 2\gamma/\Delta r_u$ Suppose that $P_{out} = P_{in}/2$; then $P_{out} = \eta_{in}/3$, like that inside a universe with radiation only. This is equivalent to say that P_{in} is due to two contributions: one coming from its radiation density content and one due to the ability of the radiation to generate matter via quantum fluctuations. The former is counterbalanced by P_{out}, the latter is the active energy excess pushing outwards the boundary. Hence the expansion of the universe is controlled by the quantum contribution of radiation fluctuation extra energy that generates matter, without which the universe would still be a radiation volume. To check this idea note that (3,3) yields $M_u^2 G/\Delta r_u = M_u c^2/2$, i.e. one half of the universe energy is equal to the first (3,17) with the proton mass replaced by that of the whole universe. The same holds for the energy density, obtained dividing both sides by V_u. So if $P_{in}/2 = \eta_{in}/3$, then $P_{out} = \eta_{in}/3$ requires $\eta_{out} = \eta_{in}/2$. Hence the right hand side yields

$$\begin{aligned} \eta_{out} &= \frac{M_u^2 G}{V_u \Delta r_u} \approx 4.2 \times 10^{-11}\,\mathrm{J/m^3} \\ P_{out} &= 2.8 \times 10^{-11}\,\mathrm{Pa} \end{aligned} \qquad (3,22)$$

as it reasonably appears comparing with η_{in} of (3,12). This result implies interesting consequences. The total contractive energy of a spherical bubble of radius Δr_u and volume V_u is $\varepsilon_\gamma = 4\pi\Delta r_u^2\gamma$. Moreover the Young-Laplace equation reads $P_{in} - P_{out} = P_{in}/2 = 2\gamma/\Delta r_u$, so that $\gamma = P_{in}\Delta r_u/4$ yields $\varepsilon_\gamma = \pi\Delta r_u^3 P_{in} = M_u c^2/2$ thanks to (3,19). Hence the whole energy of the boundary layer generating its contractive surface tension is one half of the total bulk energy of the universe, i.e. that corresponding to the net pushing effect of the big-bang quantum fluctuation only. Also, this confirms that $P_{out} = P_{in}/2$ is an external pressure opposite to P_{in} and directed towards the universe center consistently with the curvature radius Δr_u. The numerical value of γ is $\approx 6 \times 10^{15}\ \mathrm{J/m^2}$, corresponding to $M_u c^2/2A_u$. It is interesting the fact that the boundary layer can be regarded as a real matter sheet curved by the pressure difference according to the Laplace equation. The initial black hole condition (3,3) is essential for this result. Note that it is possible to write

$$\eta_{out} = \frac{3}{16\pi\Delta r_u^2}\frac{c^4}{G} = \frac{3}{4}\frac{c^4}{A_u G}; \qquad \gamma = \frac{c^4\Delta r_u}{4A_u G} \qquad (3,23)$$

i.e. the compression force at the boundary of the universe is of the order of the Planck force acting on its total surface. It is interesting to note that replacing $A_u = 4l_P^2\sigma_{\mathrm{H}}$, it is possible to express γ as a function of the Hawking entropy. Moreover, once knowing η_{out} it is easy to find the thickness of the boundary layer. This energy density is that stored in a layer surrounding the universe δr_u thick. i.e. the boundary protons and antiprotons are actually contained in a shell of volume $(4\pi/3)[(\Delta r_u + \delta r_u)^3 - \Delta r_u^3]$; so

$$V_{out} = \zeta V_u; \qquad \zeta = (1 + \delta r_u/\Delta r_u)^3 - 1 \qquad (3,24)$$

which means that in fact the size of the universe is still described just by its radius Δr_u via a correction factor ζ. Having defined η_{out} at Δr_u, it is immediate to estimate also energy, mass and number of protons/antiprotons of the boundary layer through the following equations

$$\begin{aligned} \varepsilon_{out} &= V_{out}\eta_{out} = \zeta M_u^2 G/\Delta r_u = \zeta\,\frac{c^4}{4G}\,\Delta r_u \\ m_{out} &\approx \varepsilon_{out}/c^2 = \zeta\,\frac{c^2}{4G}\,\Delta r_u \\ n_{out} &\approx m_{out}/m_p = \zeta\,\frac{c^2}{4Gm_p}\,\Delta r_u. \end{aligned}$$

If $\delta r_u \gg \Delta r_u$, then $\zeta \approx (\delta r_u/\Delta r_u)^3$; if instead $\delta r_u \approx \Delta r_u$, then $\zeta \approx 7$. Moreover, trusting to the idea that $\delta r_u \ll \Delta r_u$ at the today time Δt_u, one finds $\zeta \approx 3\delta r_u/\Delta r_u$ and then $V_{out} \approx 4\pi\Delta r_u^2\delta r_u$. Suppose that $\delta r_u \approx 10^{-15}$ m, which corresponds to the size of the proton; then $V_{out} \approx 2 \times 10^{39}\ \mathrm{m^3}$ yields $\varepsilon_{out} \approx 10^{29}$ J; i.e. the boundary layer consists of a total mass $m_{out} \approx 10^{12}$ kg, to which correspond about $n_{out} \approx 6 \times 10^{38}$ protons and antiprotons. It would be also easy with the help of (3,20) to repeat the estimates also a different past times. Going beyond the raw numerical estimates, one concludes: (i) the number density n_{out}/V_{out} is of the order of 1/3 proton per cubic meter, a figure similar to that found in V_0 of the bulk universe; (ii) the number of boundary protons results $\ll N_{in}$, as it must be according to the previous considerations; (iii) the fact that the size of the proton is of the order of one fm means that the boundary layer is actually formed by a monolayer of protons and antiprotons; also this result seems in effect quite reasonable. The connection of these conclusions with the previous (3,1), (3,6) and (3,8) will appear shortly.

Now let us explain why the presence of the proton/antiproton couples at the boundary is important for the growth of the bulk universe. Assume that the empty V_0 cells of the universe, i.e. our core vacuum, actually includes couples of virtual particles and antiparticles that annihilate and then rematerialize: whatever their specific nature might be, a simple reasoning shows that the main effect of sharing these virtual couples between bulk states and boundary states is that of transferring to the aforesaid boundary layer the properties of the bulk universe. It is essential that both virtual particles and antiparticles have equal probability of being in either state, see the next section for more details; in this sense it is possible to regard them as a couple. These forerunner quantum couples are the precursors that generate a new boundary of the universe and activate its expansion. Indeed transferring the energy early contained in any V_0 towards the boundary means reproducing at the boundary the quantum states characterizing the bulk universe, i.e. not only that of protons and antiprotons but also the vacuum energy fluctuation generating them. This also means that the universe grows by replicating part of itself outside itself; the duplication concerns of course also the virtual couples of particles and antiparticles

characterizing the core vacuum, which once more confirms why (3,12) and (3,14) have been calculated with M_u and not with m_u. So in the present model expansion does not mean merely swelling: the chance that these couples annihilate and rematerialize at the external boundary of the universe likewise as they did inside the universe, means that even the external boundary assumes the feature of the core cells V_0. In effect the previous figures recalculated with a value of δr_u slightly larger than one proton monolayer yield a proton/antiproton density comparable to that within V_0 of the bulk universe; this clearly indicates that increasing δr_u means increasing the number of boundary states allowed to protons and antiprotons. Yet proton and antiproton density in the boundary layer equal to that existing in V_0 means that the bulk of the universe has been in fact expanded by a supplementary layer $\gtrsim \delta r_u$, i.e. the actual boundary is located a step δr_u beyond the previous one, and so on by successive steps consistent with a growth rate presently given by v_u of (2,6). The driving force of this "onion growth" process is the entropy increase required by the second law: all protons and antiprotons filling the bulk universe only, anyhow distributed and arranged, would define a degree of order greater than that where some of them have the additional chance of being further arranged in the only region furthermore conceivable, i.e. that glued to the external boundary of the black hole universe. Yet the key concept is clearly the quantum uncertainty, in its most agnostic form of (2,1): being completely delocalized everywhere in the universe, the particles can preferentially be in fact wherever they ensure the most advantageous entropy and growth conditions.

However, the question that then arises immediately is: does this chance expel to the boundary exactly equal amounts of particles and antiparticles or is there preferential transfer of either kind of them? From a statistical standpoint the answer is indeed that reasonably couples of virtual particles only should share this growth mechanism: drawing randomly from a multitude of particles and antiparticles, the realistic chance is that equal numbers of either kind are involved in the quantum state change. Despite this statistical equivalence, however, the next section will emphasize why the overall effect of the entropy increment is that of increasing the matter/antimatter ratio in the bulk universe.

3.5 The problem of matter and antimatter

This section describes a mechanism really possible soon after the end of the radiation era; the couples proton/antiproton just formed from the very hot radiation field have actual physical meaning, instead of being mere statistical entities suggested by (3,12). Is useful here a reasoning similar to that of the Dirac sea, which in the present context seems physically even more appropriate than the original one: are inherent here neither infinite states occupied by electrons with negative energy nor the doubtful concept of "neutrality" conventionally defined by the presence of infinite electrons in negative energy occupied states; the Pauli principle is no longer necessary to avert a weird radiation of negative energy.

In the original Dirac idea, a photon of energy $\geq 2m_ec^2$ excites an electron in the negative state above the forbidden gap; as a result, the electron just removed appears as a standard electron that leaves behind a related positive hole, the positron. Today we know that in fact two photons of sufficient energy are able to create a couple particle/antiparticle while fulfilling the conservation laws. Let us implement here this standpoint, emphasizing however that the driving energy has now entropic character: the energetic photons necessary to modify the Dirac sea of negative energy electron states is here replaced by the entropy increase $T\delta S$ that results from the combined configuration option, bulk state and boundary state, allowed for each proton and each antiproton. The number of proton and antiproton quantum states is the large but finite N_{in}. It has been already estimated that just after the radiation era T was of the order of $10^{32} \div 10^{33}$ K; this range of values seems high enough to account for a Dirac-like process. Discuss separately what happens when one proton and one antiproton pass from their own bulk states in V_0 to their respective boundary states; two V_0 states are involved in the process, the probability that this happens is equal for both.

One proton in the first V_0 has the same energy as in the boundary state; with the proton in this latter state a hole is left behind in this V_0, i.e. a neutral vacuum state forms in the bulk universe. No constrain is necessary about the energy $T\delta S$ to allow the change from bulk to boundary state, either configuration is allowed at zero energy cost; now one V_0 state is chargeless, whereas one boundary state is positively charged.

The Dirac reasoning for an antiproton in the second V_0 sounds as follows. A proton in the negative energy state in this V_0 is excited concurrently and with the same statistical probability of the previous process; now a constrain about the excitation energy is required and reads $T\delta S \geq 2m_pc^2 + m_ec^2$. This proton is thus excited, leaves unoccupied its initial state, overcomes the forbidden gap at the right hand side and appears as an ordinary proton; a negative hole, i.e. one antiproton, results by consequence. This hole is to be regarded in the boundary state, previously raised to a positive charge state by the first proton, to ensure the local electric neutrality; the ordinary proton co-generated in the second bulk state V_0 remains inside the bulk universe together with the negative charge of one electron; this latter, necessary for the total spin conservation and for the overall bulk neutrality at the minimum energy cost, occupies the former empty vacuum state V_0 left behind from the first proton.

Clearly this mechanism requires that both a proton and an antiproton change contextually and with the same probability their bulk states, in which case we have: (i) two boundary states altogether neutral occupied by one proton and one negative antiproton, which can yield by annihilation the electromagnetic radiation trapped at the boundary of the universe and concerned since the beginning by (3,1), as confirmed by

(3,6) and (3,8); (ii) a neutral bulk state formed by one proton and one electron occupying the two volumes V_0 left vacant. Also the electroneutrality in the bulk universe is thus fulfilled thanks to the electron energy included in the energy balance of $T\delta S$. On the one hand, therefore, the equal probability of exciting statistically one proton or one antiproton is essential to ensure the neutrality of both bulk and boundary states; on the other hand, by consequence of this mechanism a couple proton/antiproton is formed in the boundary state, whereas in the bulk one proton has replaced the antiproton with the help of one electron. In fact this process removes antimatter from the bulk universe, which appears as electromagnetic radiation surrounding the universe via entropy driven process; the holes of negative energy states, i.e. protons, concurrently generated along with electrons appear as bulk matter. Moreover just the annihilation electromagnetic halo ensures the growth of the universe, which therefore does not simply swell but replicates itself far at the boundary via annihilation energy. The separation boundary-antimatter from bulk-matter was likely allowed to occur just at the beginning of the matter era, when the matter started being generated from the extremely hot radiation field consistent with its $T\delta S$. It is reasonable to think that without this separation the bulk universe would have remained in the radiation era, because the two photon mechanism previously hypothesized would have continued to produce virtual matter that however endlessly annihilated with the virtual antimatter contextually generated. Since no energy escapes from the black hole universe, $T\delta S = \delta(TS) - S\delta T$ caused decrease of internal energy and cooling of the universe, until when the temperature decrease made impossible the radiation driven formation of virtual proton/antiproton couples and the consequent antimatter expulsion to the boundary along with the concurrent formation of low T matter. Begins just now the matter era. Of course all this is possible because of the total uncertainty of the quantum particles introduced in its most agnostic form of (2,1): these particles do not need any actual travel to go from bulk to boundary of the universe, being instead totally delocalized; they are simultaneously everywhere without any chance of specify their actual location. These ideas have been exploited to discuss the EPR paradox in the frame of a relativity model entirely based on the space-time uncertainty [12].

As concerns the point (i) above, (3,6) to (3,9) and related considerations about $\hbar\omega_{n=2}$ agree with the idea that both protons and antiprotons existing at the boundary of the universe contribute with their annihilation to form the halo of electromagnetic radiation surrounding the universe.

As concerns the point (ii), the presence of the electron is evidenced simply implementing the second (3,17): the electron energy ε_e early contributed by $T\delta S$ replaces ε_p in the empty V_0 left behind by the previous proton now occupying the boundary state, so the energy density in the bulk volume V_0 becomes ε_e/V_0. To confirm this mechanism, it is enough to estimate $T = (\varepsilon_e/V_0\mathrm{a})^{1/4}$ via the black body constant $\mathrm{a} = 5.67 \times 10^{-16}\ \mathrm{J/m^3K^4}$; today's $V_0 \approx 2\mathrm{m}^3$ yields $T \approx 2.63$ K. Of course in the past, when $V_0' \ll V_0$ according to (3,20), the energy density was higher and thus the background cosmic temperature accordingly higher; the low energy of the present cosmic radiation is due to the swelling of the early V_0', formerly of the order of the Planck volume, to the size of today's V_0 that decreases the electron energy density. This conclusion agrees with the condition $n\lambda = 2\Delta r$ previously introduced to describe the evolution of the radiation field as a function of the growing universe size during the radiation era. The mechanism that originates the CBMR dates back to the early beginning of the matter era when this mechanism took place, but is operating even presently: the today wavelength, due to the swelling of the early V_0' to the current V_0, is related to the virtual couples of particles/antiparticles that feed the growth of the universe keeping constant its black hole ratio $M_u/\Delta r_u$ according to (3,20) and the concept of vacuum. The small % discrepancy from the experimental value 2.72 K of today background cosmic radiation is due to having implemented the mere rest mass of the electron, whose kinetic energy instead is presumably not exactly zero; being the electron much lighter than the proton, a relativistic correction factor in the energy balance of $T\delta S$, corresponding to $\mathrm{v}_e \approx 0.5c$ and reasonably expected, increases slightly the energy density in V_0 and allows to fit exactly the experimental value. Yet this is not the main point: the most important aim of the model is to verify the sensibleness of estimated values with respect to the available experimental data and assess the conceptual consistency of the theoretical model with the current knowledge of the universe.

4 The dark mass

A crucial point that deserves a rational explanation, hitherto not yet concerned, regards the mass $m_?$. Some comments on this mass are here reported starting from (3,9) and (3,17) and comparing the energy $\hbar\omega_n$ with $\Delta\varepsilon_u = \hbar/\Delta t_u$. One finds

$$\frac{nc\Delta t_u}{2\pi\Delta r_u} \approx \frac{n}{20}. \tag{4,1}$$

In effect, with the help of (3,1) and (3,3) the ratio at left hand side is equal to about $n\xi/2\pi$ with $\xi = 0.3$ according to (2,5) and (2,6). In section 3.1 it has been highlighted that $n = 2$ means considering electromagnetic waves surrounding the universe whose energy corresponds to the annihilation of several protons with antiprotons; also, in agreement with (3,8), for $n = 2$ the right hand side of (4,1) becomes 10^{-1}. Recall now that just a factor ten has been already found in (3,4), when describing the ratio M_u/m_u. So it seems natural to introduce this ratio into (4,1) that becomes therefore

$$\frac{M_u}{m_u}\,\frac{c\Delta t_u}{\pi\Delta r_u} \approx 1. \tag{4,2}$$

Very large numbers that fit such a simple numerical value suggest a significant physical meaning hidden in the last equa-

tion: the fact that $M_u c\Delta t_u \approx \pi m_u \Delta r_u$ is interesting because it provides a new link between M_u and m_u, i.e. according to (3,4) $m_u c^2\hbar/\Delta t_u \approx m_u c^2\hbar\omega_{n=2} + m_? c^2\hbar\omega_{n=2}$ with $\omega_{n=2} = c/\pi\Delta r_u$. Going thus beyond the mere numerical result, let us generalize (4,2) to any Δt by replacing 1 with a number $q = q(\Delta t, \Delta r)$; so the subscript u characterizing today's quantities will be omitted, whereas different values are expected for $\omega_{n=2}$ and $m_?$. Multiply both sides of the resulting equation by c^2; recalling again (3,4), (4,2) turns into

$$qmc^2\,\frac{\hbar}{\Delta t} = mc^2\hbar\omega_{n=2} + m_? c^2\hbar\omega_{n=2};\quad \omega_{n=2} = \frac{c}{\pi\Delta r}. \qquad (4,3)$$

This equation is interesting because its terms are cross linked: a couple of terms shares mc^2, another couple $\hbar\omega_{n=2}$. This shows that m and $m_?$ are correlated. Moreover the fact that this equation contains squared energy terms, brings to mind an important equation inferred in the appendix, i.e.

$$\varepsilon^2 = (pc)^2 + \varepsilon^2_{rest}. \qquad (4,4)$$

Add $\zeta mc^2\hbar/\Delta t$ to both sides of (4,3); by comparison these equations suggest the following correspondences

$$\begin{aligned} &(q+\zeta)\,mc^2\,\frac{\hbar}{\Delta t} = \varepsilon^2 \\ &mc^2\left(\hbar\omega_{n=2} + \zeta\,\frac{\hbar}{\Delta t}\right) = (pc)^2 \\ &m_? c^2\,\hbar\omega_{n=2} = \varepsilon^2_{rest} \end{aligned} \qquad (4,5)$$

being $\zeta = \zeta(\Delta t, \Delta r)$ a function of Δr and Δt whose physical meaning will appear soon. In principle these correspondences, merely based on the one-to-one association between (4,3) and (4,4) having an analogous form, propose a possible explanation of the mass $m_?$.

The universe as a whole is to be regarded like a free spinless neutral macro-particle moving at uniform speed, whose kinetic and total energy are respectively related to the terms $(pc)^2$ and ε^2; accordingly $m_?$ accounts for the rest energy of the macro-particle universe. It seems surprising that this link, suggested by mere numerical analysis of the values of $\hbar\omega_{n=2}$ and $\hbar/\Delta t_u$ of (4,2), is provided by a formula of special relativity and not of general relativity. The energies of (4,3) concern the universe as a whole and not the interaction of its parts, galaxies and stars and so on, whose gravitational dynamics is governed by the general relativity. In effect, (3,3) regards the black-hole universe as a global object, a spinless macro-particle, whose properties are due to its total mass and total size only, regardless of its complex internal structure, mass composition and mass distribution assumed homogeneous at least on large scale. A valid support to propose a rectilinear uniform motion of the whole universe comes from the fact that indeed this idea cannot be excluded by any experiment: since Galileo it is known that such an inertial motion cannot be detected by any observer inside the universe. Perhaps a harder implication of this idea could concern the hypothetical reference system R_u able to describe this motion; however also this dilemma is actually a false problem in the present model, once thinking the size of the universe as an uncertainty range $\Delta r = r_1 - r_0$ in principle similar to that introduced in section 2 to describe energy levels and angular momentum of the quantum particles. It has been emphasized: (i) that neither r_0 nor r_1 must be specified to describe the quantum properties; (ii) that in fact both coordinates are not specifiable; (iii) that this conceptual lack of information prevents specifying the reference system R_u where is defined r_0 and the actual size of Δr defined by r_1. So it is conceptually impossible, but also inessential, to specify such R_u as regards the quantum properties of a particle within the range Δr_u during the time range Δt_u: if the properties of the quantum macro-particle we call universe do not depend on r_0 or r_1 but on Δr only, then the difficulty of defining R_u, e.g. its origin, becomes marginal. Anyway, since (4,1) and (4,3) come directly from the experimental values of Δr_u and Δt_u, there is no reason to reject them; in effect (4,3) and its relativistic free particle interpretation explain why one addend concerns the mass $m_?$ and its energy $m_? c^2$ additional to the visible mass m_u of stars. Now is justified the function ζ knowing that $\varepsilon = mc^2/\sqrt{1-(\mathrm{v}/c)^2}$ and $p = m\mathrm{v}/\sqrt{1-(\mathrm{v}/c)^2}$; also these formulas are shown in the appendix in the frame of the present model. Let us rewrite the three terms of (4,4) that define the relativistic energy of the free macro-particle universe of (4,5) as a function of its displacement constant velocity v_{mp} and mass M_{mp}; this means replacing v and m with v_{mp} and M_{mp}. Hence

$$\begin{aligned} &(q+\zeta)\,mc^2\,\frac{\hbar}{\Delta t} = \frac{M^2_{mp}c^4}{1-\mathrm{v}^2_{mp}/c^2} \\ &mc^2\left(\hbar\omega_{n=2} + \zeta\,\frac{\hbar}{\Delta t}\right) = \frac{M^2_{mp}\mathrm{v}^2_{mp}c^2}{1-\mathrm{v}^2_{mp}/c^2} \\ &m_? c^2\,\hbar\omega_{n=2} = M^2_{mp}c^4. \end{aligned} \qquad (4,6)$$

Taking the ratio side by side of the first two equations one finds with the help of (3,9)

$$\frac{\mathrm{v}^2_{mp}}{c^2} = \frac{\omega_{n=2}\Delta t + \zeta}{q+\zeta};\qquad \omega_{n=2}\Delta t = \frac{c}{\pi}\,\frac{\Delta t}{\Delta r} = \frac{\mathrm{v}}{\pi c} \qquad (4,7)$$

where v is the average expansion rate of the universe at Δt. Now we impose that v_{mp} is constant via the function ζ; so

$$\mathrm{v}_{mp} = \pm c\,\sqrt{\zeta_0};\ \ \zeta = \frac{\omega_{n=2}\Delta t - \zeta_0 q}{\zeta_0 - 1};\ \ q = \frac{M}{m}\,\omega_{n=2}\Delta t \qquad (4,8)$$

i.e. q generalizes (4,2). Note that M_{mp} does not appear in these equations; it is merely defined by the third (4,6) as a function of $m_?$, on which however no hypothesis has been made. So the definitions of ζ and q hold regardless of M_{mp}. An obvious condition is $0 < \zeta_0 < 1$; moreover $q + \zeta > 0$ and $\omega_{n=2}\Delta t + \zeta > 0$ are also evident because both sides of (4,6)

and (4,7) are positive. The former condition reads $q + \zeta = (\omega_{n=2}\Delta t - q)/(\zeta_0 - 1) > 0$, the latter reads $\omega_{n=2}\Delta t + \zeta = \zeta_0(\omega_{n=2}\Delta t - q)/(\zeta_0 - 1)$; owing to the expression of q both reduce to the unique condition $1 - M/m < 0$, which is indeed true as it has been introduced since the beginning in the present quantum model. Impose also $m_?\omega_{n=2} = const$, which yields $m_?c/\pi\Delta r = const$: this equation extends (3,20) that reads $M/\Delta r = const = m/\Delta r + m_?/\Delta r$ according to (3,4). In this way M_{mp} becomes a constant. Note that owing to (3,8) this result reads $m_?c^2\hbar\sqrt{\rho G} = \varepsilon^2_{rest}$; being by definition $\rho = (m_? + m)/V$, one concludes that ε_{rest} is defined even during the early the radiation era when the visible mass was $m = 0$ and the universe volume of the order of the Planck volume V_P. So, with obvious meaning of symbols, $\varepsilon^2_{rest} = m^{(0)}_?c^2\hbar\sqrt{m^{(0)}_?G/V_P}$ and remained constant since then; hence the third (4,5) reads $m_? = m^{(0)}_?\sqrt{\rho^{(0)}/\rho}$, with $\rho^{(0)} = m^{(0)}_?/V_P$. Of course, as already noted, $1/\rho$ is an increasing function of $m_?$ because the black hole condition $M/\Delta r = const$ requires M/V decreasing function of Δr^3. In conclusion (4,5) are appropriate to describe a free macro-particle of constant mass M_{mp} moving at constant rate v_{mp}. Eventually, note that the square energies of (4,5) are actually products of two different energies, as if they would come from geometrical averages like for instance $< \varepsilon > = \pm((q + \zeta)mc^2\hbar/\Delta t)^{1/2}$. So the black hole we call universe has, as a whole, the average energy $< \varepsilon >$ of a free particle that moves with average constant momentum $< pc > = \pm(mc^2(\hbar\omega_{n=2} + \zeta\hbar/\Delta t))^{1/2}$, whereas $m_?c^2$ and $\hbar\omega_{n=2}$ defining $< \varepsilon_{rest} > = \pm(m_?c^2\hbar\omega_{n=2})^{1/2}$ appear to be the ingredients of its average rest energy. Otherwise stated, the well defined mass balance between $m_?$ and m_u proposed here appears rationally motivated: m_u is due to the capability of the universe to create ordinary visible mass after the radiation era exploiting the available big-bang fluctuation energy; the additional mass $m_?$ ensures the existence of an efficient black hole universe that does not waste uselessly its valuable energy content. So it follows also the necessity of a displacing universe. Are unavoidable at this point at least three questions: does actually the equation $< \varepsilon > = \pm\sqrt{< pc >^2 + < \varepsilon_{rest} >^2}$ admit the minus sign? could an anti-universe actually exist with a matter/antimatter mechanism equal and opposite to that described in the previous section? is our whole universe a wave/corpuscle subjected itself to the uncertainty principle?

5 Discussion

The cosmology is probably the most difficult among the physical sciences because of both its multidisciplinary conceptual basis and scarcity of experimental data, besides inferred in a limited domain of time and space consistent with the light speed: past, present and future of the whole universe must be guessed despite the space-time horizon gives us access to a limited window of observable objects only. Just for this reason the theoretical models have a special role in cosmology. Usually the experimental data validate a theoretical model; here instead seems true the exact contrary, i.e. a sound self-consistent model highlights the physical meaning of the available experimental data. In this particular context is crucial the role of quantum mechanics. The correspondence principle states that the classical physics is the limit of quantum physics for high quantum numbers, which implicitly means that just the quantum principles are the true essence of physics and thus of cosmology as well. This explains the attempt of the present model, mostly based on quantum considerations rather than on relativistic considerations. Two important experimental values, the Hubble constant and the cosmic background radiation temperature, have been estimated with accuracy enough to conclude that the physical approach of the present quantum model of the universe is basically correct. (2,1) enable the most important equations of quantum mechanics and relativity to be inferred [12,13,14,18]; their generality is also proven in particular by the ability of describing quantum fluctuations of a relativistic free particle. For instance the appendix shows how to find the well known equation $p = \mathrm{v}\varepsilon/c^2$ via $\Delta p = \mathrm{v}\Delta\varepsilon/c^2$, whose importance for the present model has been already emphasized, e.g. (2,4) and (3,7); however Δp and $\Delta\varepsilon$ are not classical ranges but quantum uncertainty ranges. So a quantum particle whose local momentum and energy are included within the respective ranges, recall the explicative results of section 2, is subjected to quantum fluctuations of p and ε that expectedly alter also its propagation rate. This fact prospects new chances for the known equations of special relativity, which here appear in fact as quantum equations subjected to the weirdness of the quantum world. Further considerations on this topic are outside the purposes of the present paper. Yet it is worth mentioning that the EPR paradox, according which particles billions of light years apart can instantaneously exchange information via the so called quantum entanglement, is explained according to the agnostic physical meaning of (2,1); the concept of distance becomes itself undetermined once disregarding the local coordinates. Renouncing even to the concept of probability density for any particle to be somewhere, replaced by the mere idea of delocalization within an uncertainty range, the concept of distance is no longer definable. So it is unphysical to expect a different quantum behavior for particles definable very close or very far apart only classically. Certainly this odd conclusion is not the only weirdness of the quantum world: as it is shown in section 2, this agnostic standpoint has unexpectedly heuristic physical meaning. One kind of weird phenomenon is the quantum fluctuation, according which any macroscopic object at rest could suddenly excited to a self-perturbed state because of a transient excess of energy, justifiable via the uncertainty principle only. The behavior of a relativistic quantum particle during a quantum fluctuation is quoted here because it is in effect pertinent to the purposes of the paper. The considerations proposed in the appendix usefully contribute to explain cosmological problems like the inflationary era. In the paper [13] it was

shown that (2,1) only are enough to infer the following corollaries: (i) equivalence of all reference systems in describing the physical laws, (ii) existence of a maximum average displacement rate allowed for any particle in its delocalization range and (iii) invariance in all reference systems of such a maximum velocity. These corollaries are in fact the basic statements of special relativity. Moreover also the equivalence principle of general relativity and the coincidence of inertial and gravitational mass were also inferred [14] along with the concept of mass as corollaries of the space-time uncertainty.

6 Appendix

This appendix sketches shortly how the relativistic momentum and energy are obtained exploiting (2,1) only; it aims to make the present paper as self-contained as possible. Let the arbitrary delocalization ranges be defined in an arbitrary reference system R, where a photon travels at speed c through $\Delta x^{(c)}$; so ((2,1)) read $\Delta x^{(c)}\Delta p_x^{(c)} = n^{(c)}\hbar = \Delta t^{(c)}\Delta\varepsilon^{(c)}$. The superscripts emphasize that the ranges are sized to fulfil the delocalization condition during an appropriate time range $\Delta t^{(c)}$. Being by definition $\Delta x^{(c)}/\Delta t^{(c)} = c$, then $c\Delta p_x^{(c)} = \Delta\varepsilon^{(c)}$. To find how the momentum and energy ranges of a massive particle traveling at rate $\mathrm{v}_x < c$ through $\Delta x^{(c)}$ scale with respect to $\Delta p_x^{(c)}$ and $\Delta\varepsilon^{(c)}$, write $\Delta x^{(c)}\Delta p_x^{(\mathrm{v})} = n^{(\mathrm{v})}\hbar = \Delta t^{(\mathrm{v})}\Delta\varepsilon^{(\mathrm{v})}$. As neither v_x nor c appear explicitly in this equation, write $n^{(\mathrm{v})}\hbar = \Delta t^{(c)}\Delta\varepsilon^{(c)} = \Delta t^{(\mathrm{v})}\Delta\varepsilon^{(\mathrm{v})}$; this is true if $\Delta t^{(c)}$ and $\Delta\varepsilon^{(c)}$ scale respectively like $\Delta t^{(\mathrm{v})} = (c/\mathrm{v}_x)\Delta t^{(c)}$, as it is reasonable, and $\Delta\varepsilon^{(\mathrm{v})} = (\mathrm{v}_x/c)\Delta\varepsilon^{(c)}$, as a consequence. Replacing these positions in the former equation, $\Delta x^{(c)}\Delta p_x^{(\mathrm{v})} = \Delta t^{(c)}(\mathrm{v}_x/c)\Delta\varepsilon^{(c)}$ yields $c\Delta p_x^{(\mathrm{v})} = (\mathrm{v}_x/c)\Delta\varepsilon^{(c)}$. Actually the superscripts can be omitted because they have been introduced for clarity of exposition only, not to identify particular range sizes; both $\Delta p_x^{(\mathrm{v})}$ and $\Delta\varepsilon^{(c)}$ are indeed completely arbitrary like v_x itself; the superscripts are also irrelevant as concerns the functional relationship between the local values of the respective variables. Hence

$$p_x = \mathrm{v}_x\varepsilon/c^2; \qquad \Delta p_x = \mathrm{v}_x\Delta\varepsilon/c^2 \tag{A1}$$

regardless of how the respective uncertainty ranges are defined. Since an identical reasoning holds in any other reference system R', one concludes that $p'_x = \mathrm{v}'_x\varepsilon'/c^2$ is an invariant of special relativity. In principle the component of velocity defining the momentum component can be positive or negative; yet squaring this equation one surely handles positive terms. So write $\varepsilon^2(\mathrm{v}_\mathrm{x}/c)^2 = (p_xc)^2$; since $\mathrm{v}_\mathrm{x}/c < 1$ for a massive particle one finds $\varepsilon^2 > (p_xc)^2$, which compels writing $\varepsilon^2 = (p_xc)^2 + \varepsilon_o^2$. Calculate the limit $p_x/\mathrm{v}_\mathrm{x}$ for $\mathrm{v} \to 0$; denoting this limit as

$$\lim_{\mathrm{v}\to 0}\frac{p_x}{\mathrm{v}_x} = m \tag{A2}$$

the concept of mass m is introduced as a consequence of the uncertainty, whereas (A1) yields $\lim_{\mathrm{v}\to 0}\varepsilon = \varepsilon_{rest} = mc^2$ in agreement with the idea that the limit must be finite; indeed no reason requires $\varepsilon \to 0$ for $\mathrm{v}_\mathrm{x} \to 0$. Thus $p_x = m\mathrm{v}_x$ is the non-relativistic form of (A1). So the previous equation yields $mc^2 = \varepsilon_o^2$, i.e.

$$\varepsilon^2 = (pc)^2 + (mc^2)^2 \tag{A3}$$

as it is well known. Hence (2,1) define themselves without additional hypotheses the concept of mass and the relativistic and non-relativistic form of the respective local variables included in the ranges Δp and $\Delta\varepsilon$. Note that merging together both equations one finds the well known expressions consistent with the Lorentz transformations. Also note that the local values of p_x and ε are exactly definable in relativity, which is substantially classical physics subjected to the covariancy principle in a four dimensional space-time context; here instead, as shown in section 2, coordinates, momentum and energy are dynamical variables random, unknown and unknowable within the respective uncertainty ranges. This is the conceptual key to understand the further considerations of this appendix. In classical physics momentum and energy of a free particle are constants; yet it is not so in the quantum world, where quantum fluctuations are allowed to occur. The crucial point is that (A1) and (A3) are quantum results, despite their form agrees of course with that of special relativity; yet, being the particles completely delocalized, the local p and ε must be intended as random values within the respective uncertainty ranges. So these equations can be accordingly handled. Let us admit that during a short time range δt even the energy of a free particle is allowed to fluctuate randomly by $\delta\varepsilon$. Since during the time transient the particle is expectedly allowed to move in a arbitrary way, (A1) is now exploited to highlight the link between $\delta\varepsilon$ and the related changes δp and $\delta\mathrm{v}$. Differentiating (A1) one finds $\delta\varepsilon = c^2\delta p/\mathrm{v} - p(c/\mathrm{v})^2\delta\mathrm{v}$: with given p and v, this result defines the functional dependence of $\delta\varepsilon$ upon arbitrary δp and $\delta\mathrm{v}$. Sum $\delta\varepsilon$ and (A1) to find $\varepsilon + \delta\varepsilon = c^2(p + \delta p)/\mathrm{v} - \varepsilon\delta\mathrm{v}/\mathrm{v}$. In general $\delta p\delta x = n\hbar$ reads $(\delta p)^2 = n\hbar\delta p/\delta x$, whereas in an analogous way $(\delta\varepsilon)^2 = n\hbar\delta\varepsilon/\delta t$. Regard just in this way $\varepsilon + \delta\varepsilon$ and $p + \delta p$; putting $\delta x = \mathrm{v}\delta t$ and replacing in the last expression to calculate $\delta(\varepsilon + \delta\varepsilon)/\delta t$, one finds

$$\begin{aligned}&(n\hbar)^{-1}(\Delta\varepsilon)^2 = (n\hbar)^{-1}(\Delta pc)^2 - \varepsilon\delta\omega\\ &\Delta\varepsilon = \varepsilon + \delta\varepsilon; \qquad \Delta p = p + \delta p.\end{aligned} \tag{A4}$$

The term $\varepsilon\delta\omega$ results because $\mathrm{v}/\delta x$ has physical dimensions of a frequency ω, so that $\delta\mathrm{v}/\delta x = \delta\omega$. As $n\hbar\omega\delta\varepsilon = \delta(\varepsilon n\hbar\omega) - \varepsilon\delta(n\hbar\omega)$, replacing this identity in the last equation one finds $(\Delta\varepsilon)^2 = (\Delta pc)^2 + n\hbar\omega\delta\varepsilon - \delta(\varepsilon n\hbar\omega)$. Let us specify this result via the position

$$n\hbar\omega = \delta\varepsilon$$

which yields also $(\Delta\varepsilon)^2 - (\Delta pc)^2 = (\delta\varepsilon)^2 - \delta(\varepsilon\delta\varepsilon)$. At left hand side appear terms containing the ranges $\varepsilon + \delta\varepsilon$ and $p + \delta p$ only, at right hand side the ranges $\delta\varepsilon$ and δp only. These latter are both arbitrary; moreover ε and p are arbitrary as well. So it

is reasonable to expect that the last equation splits into two equations linked by a constant energy ε_o

$$(\Delta\varepsilon)^2 - (\Delta pc)^2 = \varepsilon_o^2 = (\delta\varepsilon)^2 - \delta(\varepsilon\delta\varepsilon). \qquad \text{(A5)}$$

Indeed ε_o agrees with both of them just because it does not depend upon neither of them. Trivial manipulations show that the first equation yields

$$\begin{aligned} p &= \pm\frac{\varepsilon_o \mathrm{v}/c^2}{\sqrt{r_\varepsilon^2 - r_p^2(\mathrm{v}/c)^2}}; \qquad & \varepsilon &= \pm\frac{\varepsilon_o}{\sqrt{r_\varepsilon^2 - r_p^2(\mathrm{v}/c)^2}} \\ r_p &= 1 + \frac{\delta p}{p} & r_\varepsilon &= 1 + \frac{\delta\varepsilon}{\varepsilon}. \end{aligned} \qquad \text{(A6)}$$

(A5) is fulfilled even during the transient. The value of the constant ε_o is immediately found as a consequence of (A2): in agreement with (A5) $\varepsilon_o^2 = \varepsilon_{rest}^2$, because A6 hold during the time transient allowing $\delta\varepsilon$; before and after that transient one must put $\delta\varepsilon = 0$ and $\delta p = 0$ in order to have the "standard" Einstein momentum and energy of the free particle, here inferred from A1 to A3. So

$$\varepsilon_{Ein}^2 = c^2 p_{Ein}^2 + \varepsilon_{rest}^2$$

$$p_{Ein} = \pm\frac{m\mathrm{v}}{\sqrt{1-(\mathrm{v}/c)^2}}; \qquad \varepsilon_{Ein} = \pm\frac{mc^2}{\sqrt{1-(\mathrm{v}/c)^2}}.$$

It is easy now to calculate the energy and momentum gaps during the time transient δt as a function of $\delta p/p$ and $\delta\varepsilon/\varepsilon$ as follows

$$\begin{aligned} \delta l\left(\frac{m\mathrm{v}}{\sqrt{r_\varepsilon^2 - r_p^2(\mathrm{v}/c)^2}} - \frac{m\mathrm{v}}{\sqrt{1-(\mathrm{v}/c)^2}}\right) &= n_{fl}\hbar \\ \delta t\left(\frac{mc^2}{\sqrt{r_\varepsilon^2 - r_p^2(\mathrm{v}/c)^2}} - \frac{mc^2}{\sqrt{1-(\mathrm{v}/c)^2}}\right) &= n_{fl}\hbar \end{aligned} \qquad \text{(A7)}$$

where δt is the time length of the fluctuation, δl the path traveled by the particle during δt and n_{fl} the number of states allowed to the particle during the energy transient. These equations are in effect nothing else but the uncertainty equations of the fluctuation gaps $\delta p_{fl} = p_{fl} - p_{Ein}$ and $\delta\varepsilon_{fl} = \varepsilon_{fl} - \varepsilon_{Ein}$. Of course $\delta p \to 0$ and $\delta\varepsilon \to 0$ after the transient, so the amounts within parenthesis vanish, while $n_{fl} = 0$ too; i.e. the fluctuation states are no longer accessible to the particle. Taking the ratio of these expressions, one finds

$$\frac{\delta l}{\delta t} = c\,\frac{c}{\mathrm{v}}. \qquad \text{(A8)}$$

According to (A8), during a quantum fluctuation of time length δt the uncertainty range δl allowed to any quantum particle corresponds to an average displacement rate $\delta l/\delta t = c^2/\mathrm{v} > c$, i.e. *as if* the particle would really propagate at superluminal rate. The reasoning to explain this result is similar to that explaining the recession motion of celestial objects mostly as a consequence of the expansion of the space-time itself. If the fluctuation modifies the size of the energy and momentum ranges, then according to (2,1) it must modify also the space and time range sizes. Yet the space range includes all local coordinates allowed to the particle: since this latter is anywhere in the space range because it is delocalized, and not because it really travels from point to point, modifying the space size means affecting the ability of the particle of being somewhere in the universe regardless of the velocity necessary to cover the path. This explains the apparent anomaly of superluminal velocity to figure out a fluctuation driven displacement. From a mathematical point of view, this is indeed possible provided that (A7) verify two inequalities: the first is $r_p\mathrm{v}/cr_\varepsilon < 1$, to avoid imaginary quantities, the second is $r_\varepsilon^2 - r_p^2(\mathrm{v}/c)^2 < 1 - (\mathrm{v}/c)^2$, in order that both left hand sides be positive. These inequalities merge into the unique $r_\varepsilon^2 - 1 < (r_p^2 - 1)(r_\varepsilon/r_p)^2$, which yields $1 - r_\varepsilon^{-2} < 1 - r_p^{-2}$ i.e. $r_\varepsilon^{-2} > r_p^{-2}$ and thus $r_\varepsilon^2 < r_p^2$. So, being $\delta p/p > \delta\varepsilon/\varepsilon$ according to (A6), $\varepsilon/p > \delta\varepsilon/\delta p$ reads thanks to (A1) and (2,1) $\mathrm{v}/c^2 > \delta t/\delta x$ and thus $\delta x/\delta t > c^2/\mathrm{v}$ even though $\mathrm{v} < c$. (A8) is confirmed noting that it could have been obtained more quickly and easily: rewrite (2,1) as $\Delta x/\Delta t = \Delta\varepsilon/\Delta p_x$ and recall (A1) $\Delta\varepsilon/\Delta p_x = c^2/\mathrm{v}_x$; replacing the latter into the former one finds $\Delta x/\Delta t = c^2/\mathrm{v}$. This result has the same form of (A8) and (3,7); without the steps (A4) to (A8) however the properties of the quantum fluctuation would not be evident. Owing to the arbitrariness of the range sizes, nothing in principle distinguishes Δx and Δt from δl and δt; yet (A7) emphasize the specific link between δl and δt and their conjugate momentum and energy just during the quantum fluctuation. For instance, (A7) admit $r_\varepsilon = 1$ and $r_p = 1$, i.e. $\delta\varepsilon = 0$ and $\delta p = 0$, in which case $n_{fl} = 0$ because of course there are no fluctuation states; instead $\Delta p_x = 0$ and $\Delta\varepsilon = 0$ are unphysical because they deny the concept of quantum uncertainty.

In conclusion the theoretical analysis describes the effect of the extra energy transient on the space-time uncertainty of the particle during the quantum fluctuation: a massive particle can displace more than allowed by its actual velocity. Transient displacement ranges $\delta l > c\delta t$ are possible for the boundary of the universe, even though forbidden in the early Einstein derivation of momentum and energy. Indeed the relativity is substantially classical physics; yet the beauty of the theory does not admit itself quantum phenomena like the fluctuations. These phenomena are instead allowed when deriving the Einstein formulas in the quantum frame of (2,1).

It is worth emphasizing however that in the particular case $\mathrm{v} = c$ even $\delta l/\delta t$ remains always and invariably equal to c.

It is clear now that also the universe expansion is interested by these results: the previous quantum considerations, unexpected in classical relativity, help to better understand and describe the so called "inflationary era". Regard the big

bang as a vacuum fluctuation that begins at the arbitrary time t_0 and expands the primordial sphere of radius r_0 according to the concepts introduced at the beginning of section 3. During δt the displacement δl of the boundary of the universe could overcome $c\delta t$, in agreement with $\delta\varepsilon \neq 0$ and $\delta p \neq 0$. Inflation did occur when the radiation density was such that the photons were allowed moving in a medium with refractive index $n_r > 1$ and matter particles, virtual or not, were generated in the radiation field during the early beginning of the later matter era. This idea agrees with the presence at the boundary of the primordial universe of the virtual couples of particles and antiparticles generating locally via their annihilation the halo of electromagnetic radiation introduced in (3,1).

As in the present approach the ranges sizes are unknown and conceptually unknowable, it is impossible to know exactly how long lasts δt. Yet it is possible to say that after a certain time range, when $\delta\varepsilon = 0$ and $\delta p = 0$ i.e. after the end of the fluctuation, the universe expansion continued at rate compliant with the usual condition v < c.

Consider an arbitrary number of particles, assumed for simplicity non-interacting; holds for i-th of them $p_i = \mathrm{v}_i\varepsilon_i/c^2$. Let Δp and $\Delta\varepsilon$ be the momentum and energy ranges including all p_i and ε_i; being the range sizes arbitrary, it is possible to write $\Delta p = \mathrm{v}\Delta\varepsilon/c^2$ with v defined in agreement with (A1). Suppose that a quantum fluctuation starts at an arbitrary time and modifies momenta and energies of some of the particles, so that the respective ranges are modified as well; then $(c^2/\mathrm{v})\Delta p = \Delta\varepsilon$ yields $(c^2/\mathrm{v})\delta\Delta p - (c/\mathrm{v})^2\delta\mathrm{v}\Delta p = \delta\Delta\varepsilon$. Moreover (2,1), which read $\Delta\varepsilon/2\pi = n\hbar\omega$ with $\omega = 2\pi/\Delta t$ and $\Delta p/2\pi = n\hbar k$ with $k = 2\pi/\Delta x$, yield $\delta\Delta\varepsilon/2\pi = n\hbar\delta\omega$ and $\delta\Delta p/2\pi = n\hbar\delta k$. The former is the Planck equation expressed as a function of $\Delta\varepsilon/2\pi$ instead of $\Delta\varepsilon$, the latter is the De Broglie equation also expressed as a function of $\Delta p/2\pi$ instead of Δp; however being the range sizes arbitrary, unknowable and inessential as concerns the eigenvalues of the physical observables, as shown in section 2, the factor $(2\pi)^{-1}$ is trivially irrelevant. It is remarkable instead that Δx and Δt of (2,1) are regarded here as wavelength and frequency of a wave, which is in fact possible in agreement with the general character of (2,1). One finds concurrently

$$\begin{aligned} &\frac{\delta\omega}{\delta k} = \frac{\omega'}{k}; \qquad \omega' = \omega - ku; \\ &u(\omega, k) = n_r^2\,\frac{\delta\mathrm{v}}{\delta\Delta p}\,\Delta p; \qquad n_r = \frac{c}{\mathrm{v}}. \end{aligned} \tag{A9}$$

Being v and δv arbitrary, it is evident that these equations hold whatever n_r might be. This conclusion is interesting because in effect the physical meaning of these equations depends just on the features of v. Call $\mathrm{v}_p = \omega/k$ and $\mathrm{v}_g = \omega'/k$, being thus $\omega' = \omega'(k)$. For $\mathrm{v} \equiv c$ (A1) reads $\Delta\varepsilon = c\Delta p$ and describes a set of electromagnetic waves propagating in the vacuum, whence $u = 0$ i.e. $\omega'/k \equiv \omega/k \equiv c$. If $\mathrm{v} < c$ is again constant, then these equations still describe a set of light waves propagating at the same rate v_p in non-dispersive medium with refractive index n_r; yet they are also compatible with a set of massive free particles displacing at the same rate. The case where $\mathrm{v} < c$ depends on k is more interesting. The equations describe light waves propagating with different velocities in a dispersive medium dependent on n_r; the first (A9) defines the group velocity $\mathrm{v}_g \neq \mathrm{v}_p$ of the whole packet formed in the dispersive medium. Analogous conclusion holds also for the matter waves: now the displacement of matter wave packet at rate v_g is related to the maximum probability to find somewhere the set of particles; indeed the first (A9) is also obtained from $\delta(\omega'/k)/\delta k = 0$, which suggests that ω'/k corresponds to the rate with which moves the maximum of the packet defined by the dispersion curve ω'/k vs k. Both electromagnetic waves and matter particles, despite their different physical nature, are thus compatible with a unique kind of equation: their common feature is the dual wave/corpuscle nature strictly connected with the quantization condition of (2,1).

The changes $\delta\omega$ and δk have been introduced as a consequence of quantum fluctuation; in effect it would be also possible to infer from $\delta\omega = \omega\delta k/k - u\delta k$ the Einstein formula for the energy fluctuations of blackbody radiation. For brevity this point is waived here; yet, is significant the ability of the quantum fluctuation to generate packets of particle waves and packets of electromagnetic waves having similar behavior. This conclusion helps to figure out the formation of matter in the radiation field during the radiation era as superposition of electromagnetic radiation and matter wave packets, both propagating with their characteristic group velocities $\mathrm{v}_g^{(r)}$ and $\mathrm{v}_g^{(m)}$. This supports the idea of fermion/antifermion pairs formed via photon fluctuations at appropriate energy fulfilling momentum and angular momentum conservation rules. The matter waves extended to all space time available justify the presence of matter throughout the universe. Indeed it is possible to write $\delta\omega/\delta k = \delta\omega^{(r)}/\delta k^{(r)} + \delta\omega^{(m)}/\delta k^{(m)}$; then, the addends at right hand side read $\delta\omega^{(r)}/\delta k^{(r)} = \mathrm{v}_g^{(r)}$ and $\delta\omega^{(m)}/\delta k^{(m)} = \mathrm{v}_g^{(m)}$. So the extra energy transient of the fluctuation of the radiation field (term at left hand side, because $\delta\omega$ is proportional to $\delta\Delta\varepsilon$) has generated a matter wave propagating at rate in general different from that of further radiation (terms at right hand side); the quantum fluctuation of this latter could generate in turn further matter and further radiation and so on, until the available energy is sufficient to repeat the process. The matter particle propagates with a group velocity $\mathrm{v}_g^{(m)}$ having finite space length; in principle the matter wave packet can also represent a chunk of matter having finite size and given probability of being found somewhere and moving in the universe. This supports the physical meaning of (3,12) as discussed in section 3.1.

Submitted on June 19, 2013 / Accepted on June 29, 2013

References

1. Burbidge E. M., Burbidge G. R., Fowler W. A., Hoyle F. Synthesis of the Elements in Stars. *Rev. Mod. Phys.*, 1957, v. 29, 547.

2. Kiefer C. Quantum Gravity, 2nd ed. Oxford University Press, Oxford, 2007.
3. Ashtekar A. and Geroch R. Quantum theory of gravitation. *Reports on Progress in Physics*, 1974, v. 37, 1211–56.
4. Polchinski J. String Theory. Cambridge University Press, 1998.
5. Green M. B., Schwarz J. H. and Witten E. Superstring Theory. Cambridge University Press, 1987.
6. Zlatev I., Wang L., Steinhardt P. J. Quintessence, Cosmic Coincidence, and the Cosmological Constant. *Phys. Rev. Lett.*, 1999, v. 82, 896–899.
7. Lidsey J. D., Wands D., Copeland E. J. Superstring Cosmology. *Physics Report*, 2000, v. 337 (4–5), 343.
8. Kachru S., Kumar J. and Silverstein E. *Phys. Rev. D*, 1999, v. 59, 106004.
9. Harvey J./,A. String Duality and Non-supersymmetric Strings. *Phys. Rev. D*, 1999, v. 59, 026002.
10. Leonhardt A. U., Paul H. Measuring the quantum state of light. *Progr. Quant. Electr.*, 1995, v. 19, 89–130.
11. Wigner E. P. On the quantum correction for thermodynamic equilibrium. *Phys. Rev.*, 1932, v. 40, 749–759.
12. Tosto S. Spooky Action at a Distance or Action at a Spooky Distance? *Progress in Physics*, 2012, v. 1, 11–26.
13. Tosto S. Quantum Uncertainty and Relativity. *Progress in Physics*, 2012, v. 2, 58–81.
14. Tosto S. Quantum Uncertainty and Fundamental Interactions. *Progress in Physics*, 2013, v. 2, 56–81.
15. Brans C. H., Dicke R. H. Mach's Principle and a Relativistic Theory of Gravitation. *Physical Review*, 1961, v. 124 (3), 925–935.
16. Wang Y., Wang Z. Time Variation of Newton's Gravitational Constant in Superstring Theories. *Phys. Rev. Lett.*, 1986, v. 57, 1978–1981.
17. Prigogine I., Geheniau J., Gunzig E., and Nardone P. *Proc Natl Acad Sci U S A*, 1988, v. 85 (20), 7428–7432.
18. Tosto S. An analysis of states in the phase space: from quantum mechanics to general relativity. arXiv gr-qc/0807.1011.
19. Seeger P. A., Fowler W. A., Clayton D. D. Nucleosynthesis of Heavy Elements by Neutron Capture. *Astrophysical Journal Supplement*, 1965, v. 11, 121.
20. Alpher R. A., Bethe H., Gamow G. The Origin of Chemical Elements. *Physical Review*, 1948, v. 73 (7), 803–804.
21. Moffat J. W. Superluminary Universe: A Possible Solution to the Initial Value Problem in Cosmology. *Intl. J. Mod. Phys. D*, (1993), v. 2 (3), 351–65. arXiv: gr-qc/9211020.

Dynamical 3-Space: Black Holes in an Expanding Universe

David P. Rothall* and Reginald T. Cahill†

School of Chemical and Physical Sciences (CaPS), Flinders University, SA 5042, Australia
E-mail: *david.rothall@flinders.edu.au †reg.cahill@flinders.edu.au

Black holes are usually studied without including effects of the expanding universe. However in some recent studies black holes have been embedded in an expanding universe, in order to determine the interplay, if any, of these two dynamical processes. Dynamical 3-space theory contains time independent solutions for black holes, which are spatial in-flows, and separately the time dependent Hubble expansion. This theory has explained numerous puzzles in observational astrophysics and contains 3 constants; G, α - which from experimental data turns out to be the fine structure constant, and δ - which is a small but nonzero distance, possibly a Planck-type length. The Hubble expansion in the dynamical 3-space theory cannot be "switched off", forcing the study, first, of isolated black holes coexisting with the expanding universe. It is shown that a time dependent black hole and expanding universe solution exists. The nature and implications of these solutions are discussed as they evolve over time. A dynamical network of black holes and induced linking cosmic filaments forming bubble structures is discussed, as a consequence of dynamical 3-space undergoing a dynamical breakdown of homogeneity and isotropy, even in the absence of baryonic matter.

1 Introduction

The motions of stars in galaxies are strongly affected by their central massive black holes, and that of galaxies in clusters are also affected by the expansion of the universe [13]. Then the need arises to analyse black holes in the expanding universe, with the view to checking if that expansion affects black hole characteristics. There is a long history of attempts to model this phenomenon analytically; early attempts include the Einstein-Strauss model through embedding Schwarzschild black holes in the background (FLRW) universe [10], and also the well known McVittie solution [16]. This gradually lead to models (see [12] or [8] for overviews) which include the cosmological constant. The currently accepted work is based on theories of gravitation by Newton, and then extended by Hilbert and Einstein. The use of these models has generated many questions about observational phenomena, such as "supermassive" galactic central black holes [11], bore hole anomalies [1, 23], flat spiral galaxy rotation curves [20] and cosmic filaments [24]. The "dark matter" and "dark energy" parameters introduced are required in order to fit the Friedmann universe expansion equation to the type 1a supernovae [19, 22] and CMB data [14]. A more recent account of space and time [2] models time as a non-geometrical process (keeping space and time as separate phenomena), which leads to the dynamical 3-space theory. This theory is a uniquely determined generalisation of Newtonian Gravity (NG) expressed in terms of a velocity field, defined relative to observers, rather than the original gravitational acceleration field. This velocity field corresponds to a space flow, which has been detected in numerous experiments. These include gas-mode Michelson interferometer, optical fibre interferometer and coaxial cable experiments, and spacecraft Earth-flyby Doppler shift data [5]. The observational phenomena mentioned above are now gradually becoming interpreted through understanding the dynamics of space, which appears to offer an explanation for "dark matter" and "dark energy" effects [6, 7]. A brief introduction to the dynamical 3-space theory along with experimental and observational tests is given in Sections 2-5. In Sections 6 and 7 we report the discovery of exact black hole solutions embedded in an expanding universe, and discuss the nature of their evolution over time, suggesting that primordial black holes develop linking filaments, which in turn form a cosmic network with bubble structures.

2 Dynamical 3–Space

Process Physics [2] is a theory of reality which models time as a non-geometric process, with space-geometry and quantum physics being emergent and unified phenomena. The emergent geometry is thought of as a structured quantum-foam "space" and is found to be dynamic and fractal in nature, with its 3 dimensionality only approximate at micro scales. If non-trivial topological aspects of the quantum foam are ignored, it may be coarse-grain embedded in a 3-dimensional geometrical manifold. This embedding ultimately allows us to describe the dynamics of the quantum foam, or space, using a classical velocity field $\boldsymbol{v}(\boldsymbol{r}, t)$, relative to an observer with coordinate system $\boldsymbol{r}$ and t [6], and here assuming zero vorticity, $\nabla \times \boldsymbol{v} = 0$:

$$\nabla\cdot\left(\frac{\partial \boldsymbol{v}}{\partial t} + (\boldsymbol{v}\cdot\nabla)\,\boldsymbol{v}\right) + \frac{5\alpha}{4}\left((trD)^2 - tr(D^2)\right) + \\ +\delta^2\nabla^2\left((trD)^2 - tr(D^2)\right) + ... = -4\pi G\rho; \quad D_{ij} = \frac{\partial v_i}{\partial x_j} \tag{1}$$

where $\rho = \rho(\boldsymbol{r}, t)$ is the usual matter density. *

The first term involves the Euler constituent acceleration, while the $\alpha-$ and $\delta-$ terms contain higher order derivative terms and describe the self interaction of space at different scales. Laboratory, geophysical and astronomical data suggest that α is the fine structure constant $\approx 1/137$, while δ appears to be a very small but non-zero Planck-like length. The emergence of gravity arises from the unique coupling of quantum theory to the 3-space [3], which determines the "gravitational" acceleration of quantum matter as a quantum wave refraction effect,

$$\boldsymbol{g} = \frac{\partial \boldsymbol{v}}{\partial t} + (\boldsymbol{v}\cdot\nabla)\boldsymbol{v} + (\nabla\times\boldsymbol{v})\times\boldsymbol{v}_R - \frac{\boldsymbol{v}_R}{1-\dfrac{\boldsymbol{v}_R^2}{c^2}}\frac{1}{2}\frac{d}{dt}\left(\frac{\boldsymbol{v}_R^2}{c^2}\right) + \ldots \quad (2)$$

where $\boldsymbol{v}_R = \boldsymbol{v}_0 - \boldsymbol{v}$ is the velocity of matter relative to the local space. The first two terms are the Euler space acceleration, the second term explains the Lense-Thirring effect when the vorticity is non-zero, and the last term explains the precession of planetary orbits.

Neglecting relativistic effects (1) and (2) give

$$\nabla\cdot\boldsymbol{g} = -4\pi G\rho - 4\pi G\rho_{DM}, \quad (3)$$

where

$$\rho_{DM}(\boldsymbol{r}, t) \equiv \frac{5\alpha}{16\pi G}\left((trD)^2 - tr(D^2)\right) + \\ + \frac{\delta^2}{32\pi G}\nabla^2\left((trD)^2 - tr(D^2)\right). \quad (4)$$

This is Newtonian gravity, but with the extra dynamical term which has been used to define an effective 'dark matter' density. Here ρ_{DM} is purely a space/quantum foam self interaction effect, and is the matter density needed within Newtonian gravity to explain dynamical effects caused by the α and δ effects in (1). This effect has been shown to offer an explanation for the 'dark matter' effect in spiral galaxies, anomalies in laboratory G measurements, bore hole g anomalies, and the systematics of galactic black hole masses, as noted below. When $\alpha = 0$ and $\delta = 0$, (3) reduces to Newtonian gravity. The $\alpha-$term has the same order derivatives as the Euler term, and so cannot be neglected *a priori*. It was, however, missed by Newton as its consequences are not easily observable in the solar system, because of the low mass of planets relative to the massive sun. However in galaxies this term plays a major role, and the Milky Way black hole data has given evidence for that term and as well for the next higher order derivative terms.

The spatial dynamics is non-local and instantaneous, which points to the universe being highly connected, consistent with the deeper pre-space process physics. Historically this was first noticed by Newton who called it action-at-a-distance. To see this, (1) can be written as a non-linear integro-differential equation

$$\frac{\partial \boldsymbol{v}}{\partial t} = -\nabla\left(\frac{\boldsymbol{v}^2}{2}\right) - G\int d^3r'\,\frac{\rho_{DM}(\boldsymbol{r}', t) + \rho(\boldsymbol{r}', t)}{|\boldsymbol{r} - \boldsymbol{r}'|^3}(\boldsymbol{r} - \boldsymbol{r}'). \quad (5)$$

This shows a high degree of non-locality and non-linearity, and in particular that the behaviour of both ρ_{DM} and ρ manifest at a distance irrespective of the dynamics of the intervening space. This non-local behaviour is analogous to that in quantum systems and may offer a resolution to the horizon problem.

3 Evidence for the α- and δ-dynamical terms

3.1 $\delta = 0$ – early studies of dynamical 3-Space

It has been shown that dynamical 3-space flows into matter [3]. External to a spherically symmetric matter density $\rho(r)$, (1) has a time-independent radial inflow solution $v(r) \sim 1/r^{\frac{1}{2}}$ leading to the matter inward acceleration $g(r) \sim 1/r^2$. This happens because the α- and δ-dynamical terms are identically zero for this inflow speed, and explains why these significant terms were missed by Newton in explaining Kepler's Planetary Laws. However, inside a spherically symmetric mass,

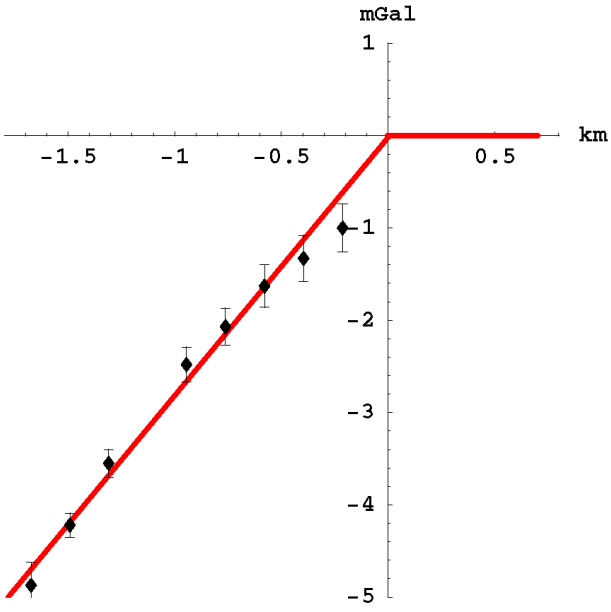

Fig. 1: The Greenland ice bore hole g anomaly data, giving $\alpha \approx 1/137$ from fitting the form in (6). The misfit at shallow depths arises from the ice not having reached the ice-shelf full density, which is a snow compactification effect. The Nevada rock bore hole data [23] also gives $\alpha \approx 1/137$. The bore hole anomaly is that gravity is stronger down a bore hole than predicted by Newtonian gravity.

*The α term in (1) has been changed by a factor of ten due to a numerical error found in the analysis of borehole data. All solutions are also altered by this factor.

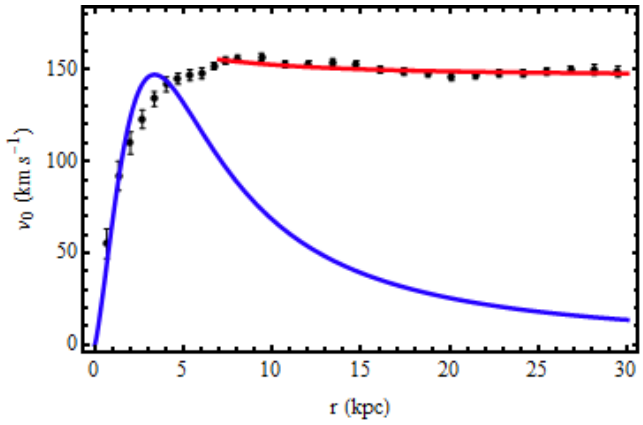

Fig. 2: The flat asymptotic star rotation speeds for the spiral galaxy NGC3198, with upper flat plot from the dynamical 3-space, while the lower form is from Newtonian gravity. The flat asymptotic form arises when $\alpha \neq 0$.

and in other circumstances, these terms play a significant dynamical role. Inside a spherically symmetric mass, such as the earth, Newtonian gravity and the new dynamics predict different matter accelerations,

$$\Delta g = g_{NG}(d) - g(d) = 20\pi\alpha G\rho d + O(\alpha^2) \qquad (6)$$

where $d < 0$ is the depth. The Greenland [1] (see Fig. 1) and Nevada bore hole data [23], reveal that $\alpha \approx 1/137$, the fine structure constant known from quantum theory. This suggests we are seeing a unification of gravity and the quantum theory.

In conventional theory black holes are required to have enormous quantities of actual in-fallen matter compressed into essentially a point-like region. Their $g \sim 1/r^2$ gravitational acceleration field is unable to explain flat spiral galaxy rotation curves, resulting in the invention of 'dark matter'. Dynamical 3-space theory however also predicts black holes in the absence of in-fallen matter, which produce a stronger acceleration field $g \sim 1/r$, as discussed below. They are spherically symmetric in-flows of space, with space not being conserved. In the absence of matter, $\rho = 0$, we set $(\boldsymbol{r}, t) = v(r)\hat{\boldsymbol{r}}$. Previous work considered solutions of (1) when $\delta = 0$, where the black hole solutions were found to have the form

$$v(r) = -\frac{\beta}{r^{\frac{5\alpha}{2}}} \qquad (7)$$

where β is an arbitrary parameter for the strength of the black hole. (1) also has straight-line filament solutions, with the form, when $\delta = 0$,

$$v(r) = -\frac{\mu}{r^{\frac{5\alpha}{4}}} \qquad (8)$$

where r is the perpendicular distance from the filament and μ is the arbitrary filament strength. The solutions (7) and (8) contain a singularity at $r = 0$ where the in-flow speed becomes infinite. Asymptotically, even when $\rho \neq 0$, these black hole solutions predict flat spiral galaxy rotation curves, for the inflow in (7) gives $g(r) = -5\alpha\beta^2/2r^{1+5\alpha} \sim -1/r$, giving the circular orbit speed $v_0(r) = (10\alpha\beta^2)^{1/2}/2r^{5\alpha/2}$, and illustrated in Fig. 2. This suggests that the 'dark matter' effect is caused by the α-dynamical term, a space self-interaction.

The Maxwell EM equations take account of the 3-space dynamics by making the change $\partial/\partial t \to \partial/\partial t + \boldsymbol{v}\cdot\nabla$. Then we obtain strong galactic light bending and lensing caused by the inflow speed in (7), or the solar light bending when $v \sim 1/r^{\frac{1}{2}}$. There are also recent direct experimental detections of the space flow velocity field by [5].

3.2 $\delta \neq 0$ – black holes and filaments

More recently the $\delta \neq 0$ scenario was considered. The form of (1) is expected as a semi-classical derivative expansion of an underlying quantum theory, where higher order derivatives are indicative of shorter length-scale physics. (1) when $\rho = 0$ has exact two-parameter, v_0 and $\kappa \geq 1$, black hole solutions

$$v(r)^2 = v_0^2(\kappa - 1)\frac{\delta}{r}\left(1 - {}_1F_1\left[-\frac{1}{2} + \frac{5\alpha}{2}, -\frac{1}{2}, -\frac{r^2}{\delta^2}\right]\right) - \\ - v_0^2\kappa\frac{8}{3}\frac{r^2}{\delta^2}\frac{\Gamma(\frac{3-5\alpha}{2})}{\Gamma(\frac{-5\alpha}{2})}\,{}_1F_1\left[1 + \frac{5\alpha}{2}, \frac{5}{2}, -\frac{r^2}{\delta^2}\right] \qquad (9)$$

where ${}_1F_1[a, b, w]$ is the confluent hypergeometric function. The parameters v_0 and κ set the strength and structure of the black hole, as discussed in [6]. (9) is a generalisation of (7), and for $r \gg \delta$ gives

$$v(r)^2 \approx A\frac{\delta}{r} + B\left(\frac{\delta}{r}\right)^{5\alpha} \qquad (10)$$

giving, from (2), $g(r) = GM(r)/r^2$, where $M(r)$ defines an "effective mass" contained within radius r, but which does not entail any actually matter,

$$M(r) = M_0 + M_0\left(\frac{r}{r_s}\right)^{1-5\alpha} \qquad (11)$$

and r_s is the distance where $M(r_s) = 2M_0$. This is shown in Fig. 3 for the Milky Way SgrA* black hole. At large r the in-flow speed becomes very slowly changing, thus predicting flat rotation curves given by [6]

$$v_{orb}(r)^2 = GM_0\left(\frac{r_s}{r}\right)^{5\alpha}\frac{1}{r_s}. \qquad (12)$$

Fig. 4 illustrates that for globular clusters and spherical galaxies the observational data implies the relationship $M_{BH} = \frac{\alpha}{2}M$. Again we see that the α-term dynamics appear to be the cause of this result, although this has yet to be derived from (1). Exact filament solutions for (1) also exist when $\delta \neq 0$, as a generalisation of (8):

$$v(r)^2 = v_0^2\frac{r^2}{\delta^2}\,{}_1F_1\left[1 + \frac{5\alpha}{4}, 2, -\frac{r^2}{2\delta^2}\right]. \qquad (13)$$

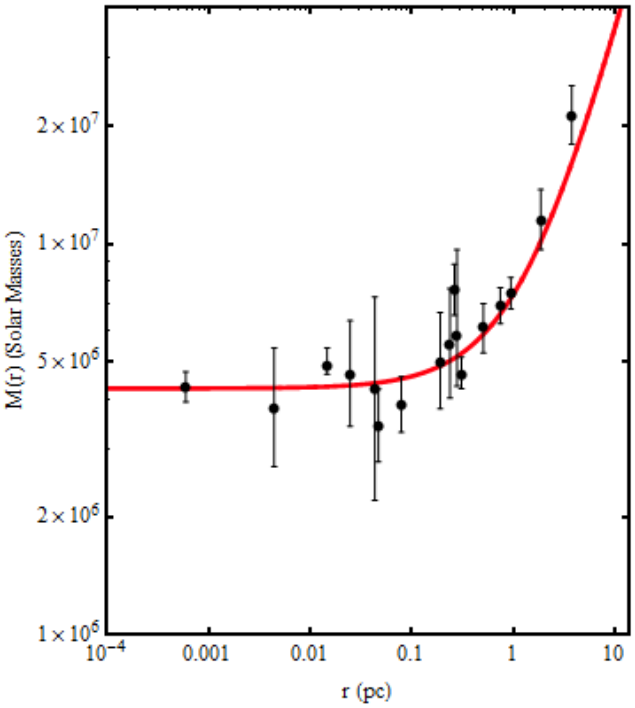

Fig. 3: Effective mass data $M(r)$ for the Milky Way SgrA* black hole, from star and gas cloud orbital data, showing the flat regime that mimics a point-like mass, but for which there is no actual matter contained within the black hole, and the linearly rising form beyond r_s =1.33pc, as predicted by (11), but which is usually attributed to a constant 'dark matter' density. This form is a direct consequence of the 3-space self-interactions in (1). The offset of the last two points indicate the presence of actual matter.

Here r is the distance perpendicular to the axis of the filament and $v(r)$ is the in-flow in that direction. The only known filament solution is for one that is infinitely long and straight. Both (9) and (13) are well behaved functions which converge to zero as $r \to 0$, i.e. the in-flow singularities are removed.

4 Expanding universe

(1) contains a time dependent expanding universe solution. Substituting the Hubble form $\boldsymbol{v}(\boldsymbol{r}, t) = H(t)\boldsymbol{r}$, and then $H(t) = \dot{a}/a$, where $a(t)$ is the universe scale factor and $\dot{a}(t) \equiv da(t)/dt$, we obtain

$$4a\ddot{a} + 10\alpha\dot{a}^2 = -\frac{16}{3}\pi G a^2 \rho \tag{14}$$

which is independent of δ. One of the key features in (14) is that even when $\rho = 0$, i.e. no matter, and $\alpha = 0$, $\ddot{a}(t) = 0$ and $a(t) = t/t_0$, and the universe is uniformly increasing in scale. Here $a(t_0) = 1$ and t_0 is the current age of the universe. This expansion of space is because the space itself is a dynamical system, and the (small) amount of actual baryonic matter merely slightly slows that expansion, as the matter dissipates space. Because of the small value of $\alpha = 1/137$, the α term only plays a significant role in extremely early epochs, but only if the space is completely homogeneous. In the limit $\rho \to 0$ we obtain the solution to (14)

$$\begin{aligned} a(t) &= \left(\frac{t}{t_0}\right)^{1/(1+5\alpha/2)} \\ H(t) &= \frac{1}{t(1+5\alpha/2)} \end{aligned} \tag{15}$$

which, as also reasoned by [17], predicts the emergence of a uniformly expanding universe after neglecting the α term. This allows a fit to the type 1a supernovae magnitude-redshift data (Fig. 5), as discussed in [7], and suggests that the dynamical 3-space theory also offers an explanation for the 'dark energy' effect. The ΛCDM parameters $\Omega_\Lambda = 0.73, \Omega_M = 0.27$, follow from either fitting to the supernovae data, or equally well, fitting to the uniformly expanding universe solution in (15) [7]. Via the dynamical 3-space solution the supernovae data gives an age for the universe of $t_0 = 13.7$ Gy.

5 Black hole – expanding universe

The Hubble solution (15) does not contain a free parameter, i.e. in the dynamical 3-space theory the universe necessarily expands, and hence it cannot be ignored when considering black holes and filaments. Since any radially flowing and time dependent $v(r, t)$ (i.e. containing both outflows and inflows) has spherical symmetry, (1) becomes, in the absence of matter

$$\begin{aligned} &\frac{\partial}{\partial t}\left(\frac{2v}{r} + v'\right) + vv'' + 2\frac{vv'}{r} + \\ &+ (v')^2 + \frac{5\alpha}{2}\left(\frac{v^2}{r^2} + \frac{2vv'}{r}\right) + \\ &+ \frac{\delta^2}{4r^4}\left(2v^2 + 2r^2(v')^2 + 6r^3v'v''\right) + \\ &+ \frac{\delta^2}{4r^4}\left(-4rvv' + 2r^2vv'' + 2r^3vv'''\right) = 0 \end{aligned} \tag{16}$$

where $v' \equiv \partial v/\partial r$. Now consider the black hole - expanding universe ansatz

$$\boldsymbol{v}(\boldsymbol{r}, t) = H(t)\boldsymbol{r} + w(r, t)\hat{\boldsymbol{r}} \tag{17}$$

where $w(r, t)$ is the spherically symmetric black hole inflow. After substituting this form we obtain a time dependent equation for $w(r, t)$. However by setting $w(r, t) = R(r)/t$ this time dependence is resolved, and (16) now may be solved for $R(r)$, implying that the Hubble outflow and black hole inflow are inseparable and compatible phenomena. Asymptotically, for $r \gg \delta$, the resulting equation for $R(r)$ has the solution

$$R(r) = -\frac{\nu}{r^{\frac{5\alpha}{2}}}, \quad \text{and so} \quad w(r, t) = -\frac{\nu}{r^{\frac{5\alpha}{2}}t} \tag{18}$$

which is the original black hole solution (7), but now with an inverse time dependence. (17) is for the black hole located at

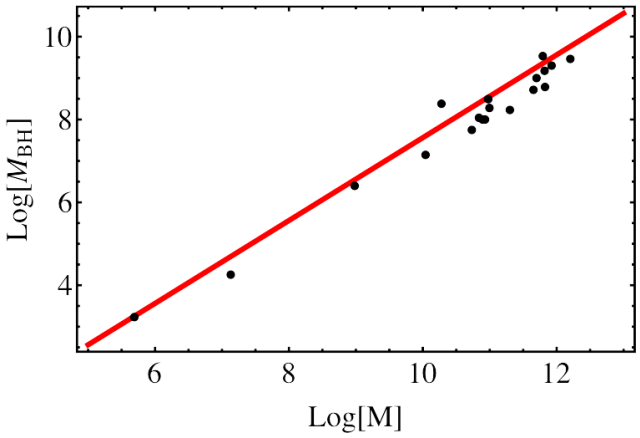

Fig. 4: Black hole masses M_{BH} vs mass M, in solar masses, for the globular clusters M15 and G1, and spherical galaxies [15]. The straight line is the relation $M_{BH} = \frac{\alpha}{2}M$, where α is the fine structure constant $\approx 1/137$. This demonstrates again the role of α in the dynamics of space and black holes.

$\boldsymbol{r} = \boldsymbol{0}$. For a black hole comoving with the local Hubble space flow the solution of (1) is

$$\boldsymbol{v}(\boldsymbol{r}, t) = H(t)\boldsymbol{r}' + w(r', t)\hat{\boldsymbol{r}}' \tag{19}$$

where $\boldsymbol{r}' = \boldsymbol{r} - a(t)\boldsymbol{r}_{BH}$ when the observer is at $\boldsymbol{r} = \boldsymbol{0}$, and the black hole is located at $a(t)\boldsymbol{r}_{BH}$. Macroscopic black holes are expected to form from coalescence of mini primordial black holes.

A consequence of (17) is that for any black hole there exists a critical radius r_c where the spatial inflow into the black hole is equal and opposite to the Hubble expansion (Fig. 6) so defining a sphere of influence. Test particles placed inside r_c are attracted to the black hole due to gravity, while those placed outside r_c, and at rest with respect to the local space, recede from it due to expansion. This critical radius is found to remain independent of time, i.e. r_c only depends on the black hole strength ν. r_c is expected to be sufficiently large that the black hole-star distance r in a galaxy today is negligible compared to r_c, i.e. $r \ll r_c$, therefore not affecting the size of the galaxies themselves. This effect would more likely be evident at a distance which galaxies are separated by, as suggested by the galaxy cluster data in [18]. For a Hubble constant $H_0 = 74$ km s^{-1} Mpc^{-1}, and using (12) for the in-flow speed, solving for $v_{orb}(r_c) = H_0 r_c$ for the Milky Way SgrA* black hole data (Fig. 3) yields $r_c = 1.6$ Mpc. For multiple black holes in the expanding space, (1) implies a more complex time evolution.

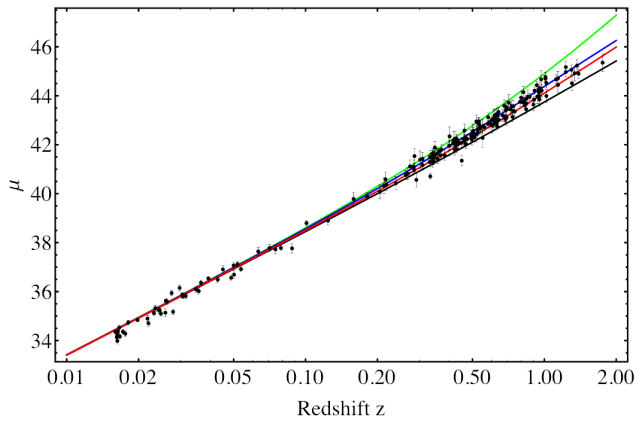

Fig. 5: Supernovae magnitude-redshift data. Upper curve (light blue) is 'dark energy' only $\Omega_\Lambda = 1$. Next curve (blue) is best fit of 'dark energy'-'dark-matter' $\Omega_\Lambda = 0.73$. Lowest curve (black) is 'dark matter' only $\Omega_\Lambda = 0$. Second lowest curve (red) is the uniformly expanding universe, and also predicted by dynamical 3-space (15).

6 Induced filaments and bubble networks

We have seen that the dynamical 3-space theory offers possible explanations for many phenomena, including that of an isolated black hole coexisting with the Hubble expansion. It also has filament solutions, in the absence of the Hubble expansion. However with multiple black holes a new feature appears to emerge, namely cosmic networks of black holes and induced filaments. First note that the black hole inflow speed in (10) is essentially very long range, resulting in the matter acceleration $g(r) \sim -1/r$, which is a key feature of these black holes, and may explain the "dark matter" effect. However this long range in-flow raises the question of how multiple black holes coexist when located within one another's sphere of influence? Fig. 7 shows the vector addition of the inflows for two black holes. This cannot be a solution of (1) as it is nonlinear and so does not have a superposition property. Whence this flow must evolve over time. Indeed the evolving flow

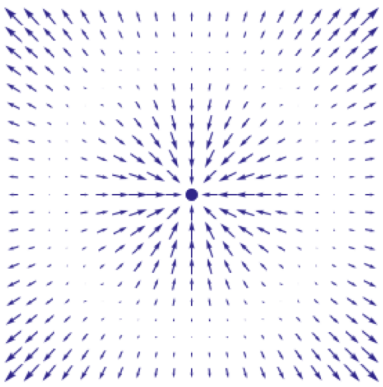

Fig. 6: Schematic 3-space velocity for an isolated black hole embedded in an expanding universe, see (17), showing radius at which flow reverses, defining the black holes sphere of influence.

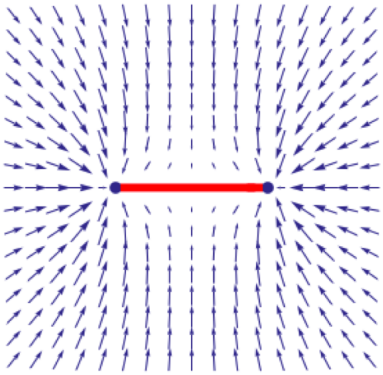

Fig. 7: 3-space in-flow velocity for two black holes located within their spheres of influence. Note the emergence of a filament forming between the black holes, indicative of a black-hole - filament network formation, see Fig. 8.

appears to form a filament connecting the two black holes. However even then there remains a long range inflow, which would lead to further filaments connecting black holes within their range of influence. These black holes are remnants of the early formation of space, and imply that (1) will undergo a dynamical breaking of symmetry, from an essentially homogeneous and isotropic 3-space, to a network of black holes and induced filaments. Note that the matter content of the universe is very small, and does not play a key role in this structure formation. A possible dynamically stable 3-space structure is shown in Fig. 8, which entails this network forming a bubble structure with the network defining a 'surface' for the bubbles. The stability of this is suggested by noting that the Hubble expansion within the interior of each bubble is now consistent with the inflow into the black holes and filaments, and so there is no longer a dynamical clash between the long range flows. Bubble structures like these are indeed found in the universe, where galaxies are observed to

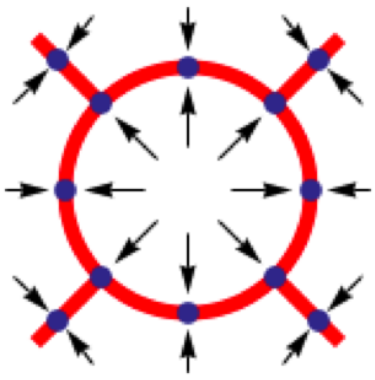

Fig. 8: 2D schematic section of a cosmic network of black holes and induced filaments. Vectors indicate 3-space flow, both within the bubble from the Hubble space expansion, and inwards to black holes (dots) and filaments (red lines). Only this bubble structure, shown here in cross-section, appears to be stable wrt the Hubble expansion.

be joined by filaments lying on spherical surfaces, filled with large voids [9, 21].

7 Conclusions

It is clear that instead of studying black-hole only cases, we need to model astrophysical and cosmological phenomena embedded in an expanding universe. The dynamical 3-space theory naturally forces us to do this, as there is no free parameter to switch off the emergent expanding universe solution, and so must be included. It has been shown that the long range black hole solutions found previously hold while embedded in an expanding universe. It is suggested that the time dependent nature of these new solutions explains in part the observed cosmic web. It appears that the dynamics of the 3-space, in the presence of primordial black holes, essentially defects in the space emerging from the quantum foam, renders a homogeneous and isotropic universe dynamically unstable, even without the presence of matter, resulting in a spatial bubble network. The long range $g \sim 1/r$ of both the black holes and induced filaments will cause matter to rapidly infall and concentrate around these spatial structures, resulting in the precocious formation of galaxies.

Submitted on July 1, 2013 / Accepted on July 6, 2013

References

1. Ander M. E. et al. *Phys. Rev. Lett.*, 1989, v. 62, 985.
2. Cahill R. T. Process physics: from information theory to quantum space and matter. Nova Science Publishers, New York, 2005.
3. Cahill R. T. Dynamical Fractal 3-Space and the Generalised Schrödinger Equation: Equivalence Principle and Vorticity Effects. *Progress in Physics*, 2006, v. 1, 27–34.
4. Cahill R. T. Dynamical 3-Space: Cosmic Filaments, Sheets and Voids. *Progress in Physics*, 2011, v. 2, 44–51.
5. Cahill R. T. Characterisation of Low Frequency Gravitational Waves from Dual RF Coaxial-Cable Detector: Fractal Textured Dynamical 3-Space. *Progress in Physics*, 2012, v. 3, 3–10.
6. Cahill R. T. and Kerrigan D. Dynamical Space: Supermassive Galactic Black Holes and Cosmic Filaments. *Progress in Physics*, 2011, v. 4, 50–64.
7. Cahill R. T., Rothall D. Discovery of Uniformly Expanding Universe. *Progress in Physics*, 2012, v. 4, 65–68.
8. Carr B. J. in Sanz J. L., Goicoechea L. J. (eds). Observational and Theoretical Aspects of Relativistic Astrophysics and Cosmology. World Scientific, Singapore, 1985.
9. De Lapparent V., Geller M. J., Huchra J. P. *ApJ*, 1986, v. 302, L1.
10. Einstein A., Straus E. G. *Rev. Mod. Phys.*, 1945, v. 17, 120.
11. Ghez A. M., Klein B. L., Morris M., Becklin E. E. *ApJ*, 1998, v. 509, 678.
12. Gibbons G. W., Maeda K. *Phys. Rev. Lett.*, 2010, v. 104, 131101.
13. Karachentsev D. et al. *A&A*, 2002, v. 383, 125.
14. Komatsu E. et al. *ApJS*, 2010, v. 192, 18.
15. Marconi A., Hunt L. K. *ApJL*, 2003, v. 589, L21.
16. McVittie G. C. *MNRAS*, 1933, v. 93, 325.
17. Melia F., Shevchuk A. S. H. *MNRAS*, 2012, v. 419, 2579.
18. Nandra R., Lasenby A. N., Hobson M. P. *MNRAS*, 2012, v. 422, 2945.

19. Perlmutter S., Aldering G., Goldhaber G., Knop R. A., Nugent P., Castro P. G., Deustua S., Fabbro S., Goobar A., Groom D. E., Hook I. M., Kim A. G., Kim M. Y., Lee J. C., Nunes N. J., Pain R., Pennypacker C. R., Quimby R., Lidman C., Ellis R. S., Irwin M., McMahon R. G., Ruiz-Lapuente P., Walton N., Schaefer B., Boyle B. J., Filippenko A. V., Matheson T., Fruchter A. S., Panagia N., Newberg H. J. M. and Couch W. J. Measurement of Ω and Λ from 42 High-Redshift Supernovae. *Astrophys.J.*, 1999, v. 517, 565–586.

20. Persic M., Salucci P., Stel F. *MNRAS*, 1996, v. 281, 27.

21. Ratcliffe A., Shanks T., Broadbent A., Parker Q. A., Watson F. G., Oates A. P., Fong R., Collins C. A. *MNRAS*, 1996, v. 281, L47

22. Riess A. G., Filippenko A. V., Challis P., Clocchiattia A., Diercks A., Garnavich P. M., Gilliland R. L., Hogan C. J., Jha S., Kirshne R. P., Leibundgut B., Phillips M. M., Reiss D., Schmidt B. P., Schommer R. A., Smith R. C., Spyromilio J., Stubbs C., Suntzeff N. B. and Tonry J. Observational Evidence from Supernovae for an Accelerating Universe and Cosmological Constant. *Astron.J.*, 1998, v. 116, 1009–1038.

23. Thomas J., Vogel P. *Phys. Rev. Lett.*, 1990, v. 65, 1173.

24. Vachaspati T. *Phys. Rev. Lett.*, 1986, v. 57, 1655.

Dilatation–Distortion Decomposition of the Ricci Tensor

Pierre A. Millette
E-mail: PierreAMillette@alumni.uottawa.ca, Ottawa, Canada

We apply a natural decomposition of tensor fields, in terms of dilatations and distortions, to the Ricci tensor. We show that this results in a separation of the field equations of General Relativity into a dilatation relation and a distortion relation. We evaluate these equations in the weak field approximation to show that the longitudinal dilatation mass relation leads to Poisson's equation for a newtonian gravitational potential, and that the transverse distortion wave relation leads to the linearized field equation of gravity in the Transverse Traceless gauge. The results obtained are in agreement with the Elastodynamics of the Spacetime Continuum.

1 Introduction

In a previous paper [1], we proposed a natural decomposition of spacetime continuum tensor fields, based on the continuum mechanical decomposition of tensors in terms of dilatations and distortions. In this paper, we apply this natural decomposition to the Ricci tensor $R^{\mu\nu}$ of General Relativity within the framework of the Elastodynamics of the Spacetime Continuum (*STCED*) [2].

2 Decomposition of the Ricci tensor

As shown in [1], the stress tensor $T^{\mu\nu}$ of General Relativity can be separated into a stress deviation tensor $t^{\mu\nu}$ and a scalar t_s according to

$$T^{\mu\nu} = t^{\mu\nu} + t_s\, g^{\mu\nu} \tag{1}$$

where

$$t^{\mu}{}_{\nu} = T^{\mu}{}_{\nu} - t_s\, \delta^{\mu}{}_{\nu} \tag{2}$$

$$t_s = \frac{1}{4}\, T^{\alpha}{}_{\alpha} = \frac{1}{4}\, T. \tag{3}$$

The Ricci curvature tensor $R^{\mu\nu}$ can also be separated into a curvature deviation tensor $r^{\mu\nu}$ (corresponding to a distortion) and a scalar r_s (corresponding to a dilatation) according to

$$R^{\mu\nu} = r^{\mu\nu} + r_s\, g^{\mu\nu} \tag{4}$$

where similarly

$$r^{\mu}{}_{\nu} = R^{\mu}{}_{\nu} - r_s\, \delta^{\mu}{}_{\nu} \tag{5}$$

$$r_s = \frac{1}{4}\, R^{\alpha}{}_{\alpha} = \frac{1}{4}\, R \tag{6}$$

where R is the contracted Ricci curvature tensor.

Using (1) to (6) into the field equations of General Relativity [3, see p. 72],

$$R^{\mu\nu} - \frac{1}{2}\, g^{\mu\nu} R = -\kappa T^{\mu\nu} \tag{7}$$

where $\kappa = 8\pi G/c^4$ and G is the gravitational constant, we obtain a separation of the field equations of General Relativity into dilatation and distortion relations respectively:

$$\begin{aligned} &\text{dilatation} : r_s = -\kappa t_s \\ &\text{distortion} : r^{\mu\nu} = \kappa t^{\mu\nu}. \end{aligned} \tag{8}$$

The dilatation relation of (8) can also be expressed as

$$R = -\kappa T. \tag{9}$$

The distortion-dilatation separation of tensor fields is thus also applicable to the field equations of General Relativity, resulting in separated dilatation and distortion relations. This result follows from the geometry of the spacetime continuum (*STC*) used in General Relativity being generated by the combination of all deformations present in the *STC* [2].

3 Weak field approximation

We evaluate these separated field equations (8) in the weak field approximation to show that these relations satisfy the massive longitudinal dilatation and massless transverse distortion results of *STCED* [2].

In the weak field approximation [4, see pp. 435–441], the metric tensor $g_{\mu\nu}$ is written as $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$ where $\eta_{\mu\nu}$ is the flat spacetime diagonal metric with signature $(-\ +\ +\ +)$ and $|h_{\mu\nu}| \ll 1$. The connection coefficients are then given by

$$\Gamma^{\mu}{}_{\alpha\beta} = \frac{1}{2}\eta^{\mu\nu}(h_{\alpha\nu,\beta} + h_{\beta\nu,\alpha} - h_{\alpha\beta,\nu}) \tag{10}$$

or, after raising the indices,

$$\Gamma^{\mu}{}_{\alpha\beta} = \frac{1}{2}(h_{\alpha}{}^{\mu}{}_{,\beta} + h_{\beta}{}^{\mu}{}_{,\alpha} - h_{\alpha\beta}{}^{,\mu}). \tag{11}$$

The Ricci tensor is also linearized to give

$$R_{\mu\nu} = \Gamma^{\alpha}{}_{\mu\nu,\alpha} - \Gamma^{\alpha}{}_{\mu\alpha,\nu} \tag{12}$$

which becomes

$$R_{\mu\nu} = \frac{1}{2}(h_{\mu}{}^{\alpha}{}_{,\nu\alpha} + h_{\nu}{}^{\alpha}{}_{,\mu\alpha} - h_{\mu\nu,\alpha}{}^{\alpha} - h^{\alpha}{}_{\alpha,\mu\nu}). \tag{13}$$

The contracted Ricci tensor

$$R = g^{\mu\nu} R_{\mu\nu} \simeq \eta^{\mu\nu} R_{\mu\nu} \tag{14}$$

then becomes

$$R = \frac{1}{2}\eta^{\mu\nu}(h_{\mu}{}^{\alpha}{}_{,\nu\alpha} + h_{\nu}{}^{\alpha}{}_{,\mu\alpha} - h_{\mu\nu,\alpha}{}^{\alpha} - h^{\alpha}{}_{\alpha,\mu\nu}) \tag{15}$$

which, after raising the indices and re-arranging the dummy indices, simplifies to

$$R = h^{\alpha\beta}{}_{,\alpha\beta} - h^{\alpha}{}_{\alpha,\beta}{}^{\beta}. \tag{16}$$

4 Dilatation (mass) relation

Making use of (16) and (6) into the dilatation relation (9), we obtain the *longitudinal dilatation mass relation*

$$h^{\alpha}{}_{\alpha,\beta}{}^{\beta} - h^{\alpha\beta}{}_{,\alpha\beta} = \kappa T \tag{17}$$

and, substituting for κ from (7) and $T = \rho c^2$ from (30) of [2],

$$\nabla^2 h^{\alpha}{}_{\alpha} - \partial_\alpha \partial_\beta h^{\alpha\beta} = \frac{8\pi G}{c^2}\rho \tag{18}$$

where ρ is the rest-mass density. This equation is shown to lead to Poisson's equation for a newtonian gravitational potential in the next section.

The second term of (18) would typically be set equal to zero using a gauge condition analogous to the Lorentz gauge [4, see p. 438]. However, the second term is a divergence term, and it should not be set equal to zero in the general case where sources may be present.

4.1 Static newtonian gravitational field

We consider the metric perturbation [4, see pp. 412–416]

$$\begin{aligned} h_{00} &= -2\Phi/c^2 \\ h_{ii} &= 0, \quad \text{for } i = 1, 2, 3 \end{aligned} \tag{19}$$

where Φ is a static (i.e. time independent) newtonian gravitational field. Then the term

$$h^{\alpha\beta}{}_{,\alpha\beta} = h^{00}{}_{,00} = 0 \tag{20}$$

and (17) becomes

$$\nabla^2 h^{0}{}_{0} = \kappa T. \tag{21}$$

Using h_{00} from (19) and κ from (7), (21) becomes

$$\nabla^2 \Phi = \frac{4\pi G}{c^2} T. \tag{22}$$

Substituting for $T = \rho c^2$ from (30) of [2], we obtain

$$\nabla^2 \Phi = 4\pi G \rho \tag{23}$$

where ρ is the mass density. This equation is Poisson's equation for a newtonian gravitational potential.

5 Distortion (wave) relation

Combining (13) and (16) with (5) and (6) into the distortion relation of (8), we obtain the *transverse distortion wave relation*

$$\begin{aligned} &\frac{1}{2}\left(h_{\mu\alpha,\nu}{}^{\alpha} + h_{\nu\alpha,\mu}{}^{\alpha} - h_{\mu\nu,\alpha}{}^{\alpha} - h^{\alpha}{}_{\alpha,\mu\nu}\right) - \\ &-\frac{1}{4}\eta_{\mu\nu}\left(h^{\alpha\beta}{}_{,\alpha\beta} - h^{\alpha}{}_{\alpha,\beta}{}^{\beta}\right) = \kappa t_{\mu\nu} \end{aligned} \tag{24}$$

where $t_{\mu\nu}$ is obtained from (2) and (3). This equation can be shown to be equivalent to the equation derived by Misner *et al* [4, see their Eq.(18.5)] from which they derive their linearized field equation and transverse wave equation in the Transverse Traceless gauge [4, see pp. 946–950]. This shows that this equation of the linearized theory of gravity corresponds to a transverse wave equation.

This result highlights the importance of carefully selecting the gauge transformation used to simplify calculations. For example, the use of the Transverse Traceless gauge eliminates massive solutions which, as shown above and in [2], are longitudinal in nature, while yielding only non-massive (transverse) solutions for which the trace equals zero.

6 Discussion and conclusion

In this paper, we have applied a natural decomposition of tensor fields, in terms of dilatations and distortions, to the Ricci tensor. We have shown that this results in a separation of the field equations of General Relativity into a dilatation relation and a distortion relation. We have evaluated these equations in the weak field approximation to show that the longitudinal dilatation mass relation leads to Poisson's equation for a newtonian gravitational potential, and that the transverse distortion wave relation leads to the linearized field equation of gravity in the Transverse Traceless gauge. The results obtained are thus found to be in accord with the Elastodynamics of the Spacetime Continuum.

Acknowledgements

The author thanks Dr. Dmitri Rabounski for his editorial support of the *STCED* research.

Submitted on July 2, 2013 / Accepted on July 5, 2013

References

1. Millette P. A. On the Decomposition of the Spacetime Metric Tensor and of Tensor Fields in Strained Spacetime. *Progress in Physics*, 2012, v. 4, 5–8.
2. Millette P. A. Elastodynamics of the Spacetime Continuum. *The Abraham Zelmanov Journal*, 2012, v. 5, 221–277.
3. Wald R. M. General Relativity. The University of Chicago Press, Chicago, 1984.
4. Misner C. W., Thorne K. S., Wheeler J. A. Gravitation. W. H. Freeman and Co., San Francisco, 1973.

The Dark Side Revealed: A Complete Relativity Theory Predicts the Content of the Universe

Ramzi Suleiman

University of Haifa, Haifa 31509, Israel. E-mail: suleiman@psy.haifa.ac.il

Dark energy and dark matter constitute about 95% of the Universe. Nonetheless, not much is known about them. Existing theories, including General Relativity, fail to provide plausible definitions of the two entities, or to predict their amounts in the Universe. The present paper proposes a new special relativity theory, called *Complete Relativity theory (CR)* that is anchored in Galileo's relativity, but without the notion of a preferred frame. The theory results are consistent with Newtonian and Quantum mechanics. More importantly, the theory yields natural definitions of dark energy and dark matter and predicts the content of the Universe with high accuracy.

1 Introduction

1.1 Dark energy

The nature of dark energy ranks among the very most compelling of all outstanding problems in physical science [1, 2]. Conclusive evidence from supernovas and other observations show that, despite gravitation, the Universe is expanding with acceleration [3–6]. No existing theory is capable of explaining what dark energy is, but it is widely believed that it is some unknown substance with an enormous antigravitational force, which drives the galaxies of our Universe apart. It is also well established that at our time the Universe is comprised of $\approx 4.6\%$ atoms, $\approx 72\%$ dark energy and $\approx 23\%$ dark matter (see e.g., [1]). One explanation for dark energy is founded on Einstein's Cosmological Constant (λ), despite the fact that Einstein himself abandoned his constant, calling it his biggest mistake. According to this explanation the Universe is permeated by an energy density, constant in time and uniform in space. The big problem with this explanation is that for $\lambda \neq 0$ it requires that the magnitude of λ be $\approx 10^{120}$ (!) times the measured ratio of pressure to energy density [1].

An alternative explanation argues that dark energy is an unknown dynamical fluid, i.e., one with a state equation that is dynamic in time. This type of explanation is represented by theories and models which differ in their assumptions regarding the nature of the state equation dynamics [7–9]. This explanation is no less problematic since it entails the prediction of new particles with masses thirty-five orders of magnitude smaller than the electron mass, which might imply the existence of new forces in addition to gravity and electromagnetism [1]. At present there is no persuasive theoretical explanation for the existence, dynamics and magnitude of dark energy and its resulting acceleration of the Universe.

1.2 Dark matter

Dark matter is more of an enigma than dark energy. Scientists are more certain about what dark matter is not, than about what it is. Some contend that it could be Baryonic matter tied up in brown dwarfs or in chunks of massive compact halo objects "or MACHOs" [10, 11], but the common prejudice is that dark matter is not baryonic, and that it is comprised of particles that are not part of the "standard model" of particle physics. Candidates that were considered include very light axions and Weakly Interacting Massive Particles (WIMPs) which are believed to constitute a major fraction of the Universe's dark matter [2, 12–14].

Given the frustrating lack of knowledge about the nature of dark energy and dark matter, most experts contend that understanding the content of the Universe and its cosmic acceleration requires nothing less than "discovering a new physics" [14]. As example, the Dark Energy Task Force (DETF), summarized its 2006 comprehensive report on dark energy by stating that there is consensus among most physicists that "nothing short of a revolution in our understanding of fundamental physics will be required to achieve a full understanding of the cosmic acceleration" [1, see p. 6]. This statement includes the possibility of reconsidering Einstein's Special and General Relativity altogether.

The present paper meets the challenge by proposing a new relativity theory. The proposed theory, *which I term Complete Relativity Theory (or CR)*, is anchored in Galileo's relativity, but without the notion of a preferred frame. Alternatively, the theory could be seen as a generalization of the Doppler Formula [15, 16] to account for the relative dynamics of moving objects of mass. The theory's results are consistent with Newtonian mechanics and with Quantum mechanics. More importantly, the theory yields relativistic definitions of dark energy and dark matter, describes their dynamics and predicts the content of the Universe with impressive accuracy.

The following sections describe the theory for the special case of zero forces, resulting in constant relative velocities. I derive its time, distance, density, and energy transformations (sections 2.1–2.3) and compare the derived energy-term with Newton's and Einstein's *Special Relativity* terms. Section 3, which constitutes the core of this paper, puts forward a relativistic definition of dark energy and dark matter, describes

their dynamics as function of the relative velocity $\beta = \mathrm{v}/\mathrm{c}$, and calculates the present content of the Universe. Section 4 concludes with a brief discussion.

2 Complete Relativity (*CR*) theory postulates and transformations

CR theory rests on two postulates:

1. The magnitudes of ***all*** physical entities, as measured by an observer, depend on the relative motion of the observer with respect to the rest frame of the measured entities.
2. All translations of information from one frame of reference to another are carried by light or electromagnetic waves of equal velocity.

Note that postulate 1 applies to all measured entities, including the velocity of light. Thus, *CR* treats the velocity of light as a relativistic quantity and not as an invariant one as postulated by Einstein's *SR*.

2.1 Time transformation

The derivation of the time transformation of *CR* is similar to the derivation of the Doppler Formula, except that *CR* treats the relative time of a moving object with constant velocity, instead of the frequency of a traveling wave.

Consider the two frames of reference F and F' shown in Figure 1. Assume that the two frames are moving away from each other at a constant velocity v. Assume further that at time t_1 in F (and t'_1 in F') a body starts moving in the $+x$ direction from point x_1 (x'_1 in F') to point x_2 (x'_2 in F'), and that its arrival is signaled by a light pulse, which emits exactly when the body arrives at its destination. Denote the times of arrivals in F and F' by t_2 and t'_2, respectively. Finally, assume that the start times in F and F' are synchronized. Without loss of generality, we can set $t_1 = t' = 0$ and $x_1 = x'_1 = 0$.

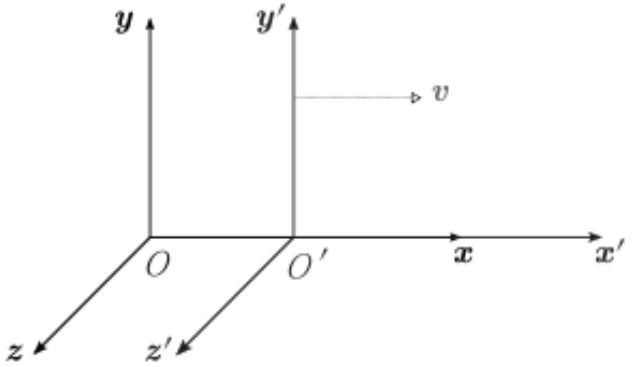

Fig. 1: Two observers in two reference frames moving with velocity v with respect to each other

The end time t_2, measured in F, equals the end time t'_2 plus the time δt which takes the light beam signaling the body's arrival at x_2 to reach the observer in F, or: $t_2 = t'_2 + \delta t$. But $\delta t = d/c$ where d is the distance (measured in F) travelled by F' relative to F, and ***c* is the velocity of light as measured in *F***. But $d = vt_2$, thus we can write:

$$t_2 = t'_2 + \frac{vt_2}{c} = t'_2 + \beta t_2 \,, \tag{1}$$

where $\beta = \frac{v}{c}$. Defining $t_2 = t$, $t_2 = t'$ and $\hat{t} = t/t'$, we get:

$$\hat{t} = \frac{t}{t'} = \frac{1}{1-\beta} \,. \tag{2}$$

Equation (2) is identical to the Doppler Formula, except that the Doppler Effect describes red- and blue-shifts of waves propagating from a departing or approaching wave source, whereas the result above describes the time transformation of moving objects. Note that $1/(1-\beta)$ is *positive* if F and F' *depart* from each other, and *negative* if they *approach* each other.

For the *round trip* from F and back, synchronization of the start time is not required. For this case the total relative time is given by (See Appendix, section1):

$$\hat{t} = \frac{t}{t'} = \frac{2}{1-\beta^2} \,. \tag{3}$$

For the one-way trip and a *departing* F' at velocity β ($0 \leqslant \beta \leqslant 1$), the proposed theory (*CR*) and Einstein's Special Relativity (*SR*) yield similar predictions, although the time dilation predicted by *CR* is larger than that predicted by *SR* (see Fig. 1Aa in the Appendix). Conversely, for an *approaching* F' ($\beta < 0$), *CR* predicts that the internal time measured at F will be *shorter* than that measured at F'. For the round trip the results of *CR* and *SR* (in $-1 \leqslant \beta \leqslant 1$) are qualitatively similar, except that the time dilation predicted by *CR* is larger than that predicted by *SR* (see Fig. 1Ab in the Appendix). For small β values the two theories yield almost identical results.

Note that the assumption that information is translated by light should not be considered a limitation of the theory, since its results are directly applicable to physical systems which use different transporters of information between two reference frames.

2.2 Distance transformation

The time duration, in frame F, of the event described above is equal to:

$$t_2 = \frac{x_2 - x_1}{c} = \frac{x_2}{c} \,, \tag{4}$$

where c is the velocity of light as measured in F. Similarly, the time duration of the event in F' could be written as:

$$t'_2 = \frac{x'_2 - x'_1}{c'} = \frac{x'_2}{c'} \,, \tag{5}$$

where c' is the velocity of light as measured in F'. From equations (4) and (5) we obtain:

$$\frac{x_2}{x'_2} = \frac{c'}{c}\frac{t_2}{t'_2} = \frac{c+v}{c}\frac{t_2}{t'_2} = \left(1 + \frac{v}{c}\right)\frac{t_2}{t'_2} = (1+\beta)\frac{t_2}{t'_2} \,. \tag{6}$$

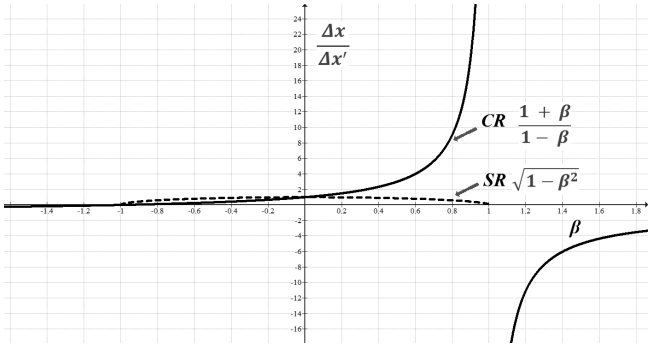

Fig. 2: Distance transformation.

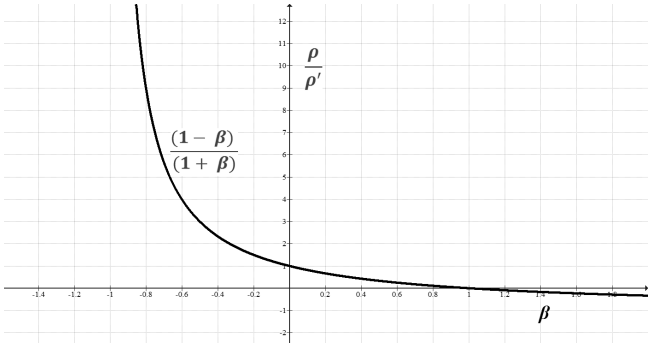

Fig. 3a: Density.

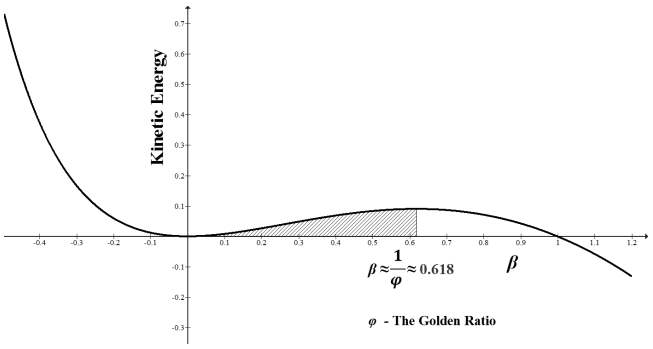

Fig. 3b: Energy.

Fig. 3: Density and energy as functions of velocity.

Substituting t_2/t_2' from (2) in (6) and denoting $x_2 = x$, $x_2' = x'$ and $\hat{x} = x_2/x_2'$ we get:

$$\hat{x} = \frac{x_2}{x_2'} = \frac{1+\beta}{1-\beta}. \tag{7}$$

The relative distance $\hat{x} = \Delta x \,/\, \Delta x' = (x_2 - x_1) \,/\, (x_2' - x_1')$ as a function of β, together with the respective relative distance according to *SR* (in dashed black) are shown in Figure 2. As shown by the figure, while *SR* prescribes that irrespective of direction, objects moving relative to an internal frame will contract, *CR* predicts that a moving object will contract or expand, depending on whether it approaches the internal frame or departs from it. For relative velocities exceeding the velocity of light ($\beta > 1$), *CR* predicts that $\hat{x}$ will become negative. Since $\Delta x'$ is positive, this implies that for bodies departing from an internal frame with a velocity higher than the velocity of light, the length of a rod of rest-length l_0, placed along the x axis, will be negative.

2.3 Density and energy transformations

Similar analyses for the density and kinetic energy (see Appendix, section 2) yield the following transformations:

Density:

$$\hat{\rho} = \frac{\rho}{\rho'} = \frac{(1-\beta)}{(1+\beta)} \tag{8}$$

and energy:

$$E = \frac{1}{2} m_0 c^2 \beta^2 \frac{(1-\beta)}{(1+\beta)}, \tag{9}$$

where m_0 is the rest mass in F'. Note that for $\beta \to 0$ (or $v \ll c$) *CR* reduces to Newton's mechanics ($\hat{t} = \hat{x} = \hat{\rho} = 1$, $E = \frac{1}{2} m v^2$). Figures 3 (a & b) depict the density and energy as functions of the velocity β. As shown by the figure the density of departing bodies relative to an observer in F is predicted to *decrease* with β, reaching zero for velocity equaling the speed of light. For bodies approaching the observer ($\beta < 1$) *CR*, similar to *SR*, predicts that the relative density will increase nonlinearly, from $\rho = \rho' = \rho_0$ at $\beta = 0$, to infinitely high values as β approaches -1. For $\beta < -1$ and $\beta > 1$, *CR* predicts that the relative density, as measured in F, will be negative.

The kinetic energy displays a non-monotonic behavior with two maxima: one at negative β values (approaching bodies) and the other at positive β values (departing bodies). The points of maxima (see Appendix, section 2) are $\beta_1 = \varphi - 1 \approx \approx 0.618$, and $\beta_2 = -\varphi \approx -1.618$, where φ is the Golden Ratio defined as $\varphi = \frac{\sqrt{5}+1}{2} \approx 1.618$ (see derivation in Appendix, section 2). The predicted decline in kinetic energy at velocities above $\beta \approx 0.618$ (see Fig. 3b), despite the decrease in velocity, suggests that mass and energy transform gradually from normal mass and energy to unobservable (dark) mass and energy.

The maximal kinetic energy at $\beta \approx 0.618$ is equal to:

$$\begin{aligned} E_{max} &= \frac{1}{2} m_0 c^2 (\varphi - 1)^2 \frac{1 - (\varphi - 1)}{1 + (\varphi - 1)} = \\ &= \frac{1}{2} m_0 c^2 (\varphi - 1)^2 \frac{(2 - \varphi)}{\varphi}. \end{aligned} \tag{10}$$

Since $\varphi - 1 = \frac{1}{\varphi}$ (See Appendix, section 2), Eq. 10 could be rewritten as:

$$E_{max} = \frac{1}{2} m_0 c^2 \frac{(2-\varphi)}{\varphi^3}. \tag{11}$$

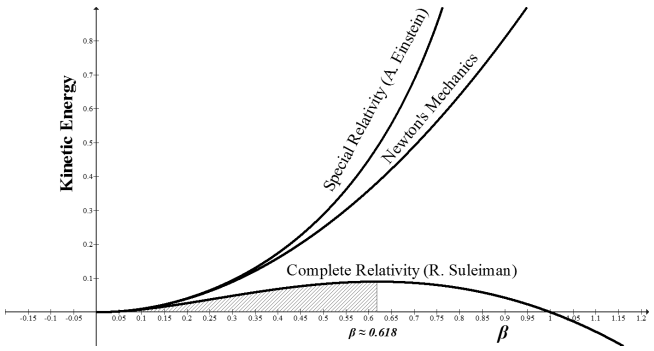

Fig. 4: Energy as a function of velocity according to three theories.

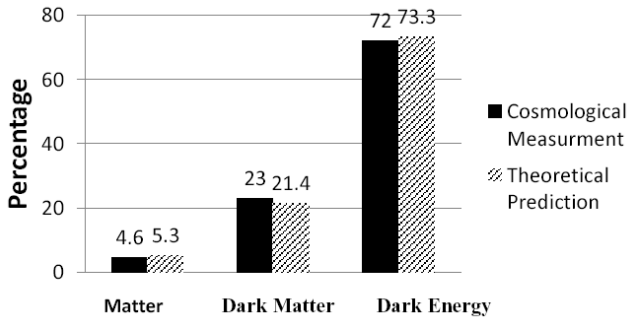

Fig. 5: Comparison between *CR*'s prediction of the content of the Universe and cosmological measurements

Substituting $\varphi = \frac{\sqrt{5}+1}{2}$ we obtain:

$$E_{max} \approx 0.04508497 m_0 c^2. \quad (12)$$

Notably, the energy-mass equivalent according to Eq. 12 is only $\approx 4.51\%$ of the amount predicted by the Einstein's famous equation $E = mc^2$. The above result is consistent with cosmological findings indicating that the percentage of Baryonic matter in the Universe is $\approx 4.6\%$. No less important the mass/energy conversion ratio (≈ 0.04508497) is precisely half of L. Hardy's probability of entanglement (0.09016994) [17–19]. This result confirms with a recent experimental finding [20], which demonstrated that applying a magnetic field at right angles to an aligned chain of cobalt niobate atoms, makes the cobalt enter a quantum critical state, in which the ratio between the frequencies of the first two notes of the resonance equals the Golden Ratio; the highest-order *E8* symmetry group discovered in mathematics [21].

For positive β values (departing objects) Figure 4 depicts *CR*'s energy function $E(\beta)$ together with the energy terms of Newton and Einstein's Special Relativity. As could be seen, while the latter theories predict that energy is strictly increasing with velocity, *CR* predicts a non-monotonic relationship with a maximum at $\beta \approx 0.618$ (the Golden ratio). As I shall show in the following section, this non-monotonic nature holds the key for explaining dark matter and dark energy.

3 The content of the Universe

The energy function Eq. 9 suggests that dark energy at a given velocity could be interpreted as the *difference between the energy measured at the internal frame and the energy measured at the external frame*. In other words, ***dark energy is defined as the energy loss due to relativity***. In formal terms, denote the energy at the internal and external frames by E' and E respectively, the kinetic energy measured at the internal and external frames could be expressed as: $E(\beta) = \frac{1}{2} m_0 c^2 \beta^2$ and $E'(\beta) = \frac{1}{2} m_0 c^2 \beta^2 \frac{(1-\beta)}{(1+\beta)}$, respectively, and the amount of dark energy, $DE(\beta)$, could be expressed as:

$$DE(\beta) = E'(\beta) - E(\beta) = \\ = \frac{1}{2} m_0 c^2 \beta^2 \left(1 - \frac{1-\beta}{1+\beta}\right) = m_0 c^2 \frac{\beta^3}{1+\beta}. \quad (13)$$

Similarly, ***dark matter, $m\_(\beta)$, at a given velocity is defined as the relativistic loss of matter at that velocity***. In other words, it equals the difference between the mass of normal matter measured at the internal and external frames. In formal notation: $m\_(\beta) = m_0 - m(\beta)$. Using the density transformation (Eq. 13), dark matter, $m(\beta)$, could be expressed as:

$$m\_(\beta) = m_0 - m(\beta) = m_0 \left(1 - \frac{1-\beta}{1+\beta}\right) = m_0 \left(\frac{2\beta}{1+\beta}\right). \quad (14)$$

The standard cosmological model of the Universe prescribes that it is comprised mainly of dark energy and dark matter (around 72% and 23%, respectively), with only less than 5% normal (Baryonic) matter. To compare matter with energy I use the matter-energy equivalence depicted in Eq. 12, according to which every unit of mass is equivalent to $\approx 0.045\, c^2$ energy units. Figure 5 depicts the dynamics of normal matter, dark matter, and dark energy as functions of β in the range $0 \leqslant \beta \leqslant 1$. Calculating the percentage of each component at $\beta = \varphi - 1 \approx 0.618$, or equivalently at redshift $z \approx 0.382$ (see Appendix, section 3) (yields $\approx 5.3\%$ Baryonic matter, $\approx 21.4\%$ dark matter, and $\approx 73.3\%$ dark energy, which is in excellent fit with current cosmological observations (See Fig. 6).

Statistical comparisons between the empirical and theoretical distributions of matter, dark matter, and dark energy, show that the difference is not significant ($p > 0.699$, Kolmogorov-Smirnov test). For velocities higher than $\beta = \frac{v}{c} \approx 0.618$ we get slightly different compositions. For example, for $\beta = 0.9$ (redshift $z \approx 0.474$) we get $\approx 89.4\%$ dark energy, $\approx 10\%$ dark matter and $\approx 0.6\%$ Baryonic matter. The average proportions in the range $0 \leqslant \beta \leqslant 1$ are about 85.80% dark energy, 12.35% dark matter and 1.85% Baryonic matter.

4 Concluding remarks

The biggest challenge of standard cosmology nowadays is to find a natural and more fundamental way to explain the detected presence of dark energy and dark matter. Most physicists agree that if this challenge is not met in the near future, then nothing less than "discovering a new physics" [14] and "a revolution in our understanding of fundamental physics" [2] will be required.

The present paper responds to the challenge by proposing a new relativity theory that is based on Galileo's relativity, but without the notion of a preferred frame. The analyses reveal that for low velocities the theory confirms with Newtonian mechanics and for high velocities it confirms with main predictions of quantum mechanics. More important for the present context, the proposed theory puts forward, for the first time, plausible definitions of dark matter and dark energy. The two entities are defined simply as the unobserved (dark) side of the matter-energy in the Universe. This definition yields formal expressions for the two entities which enable to predict the present content of the Universe with high accuracy. Two additional important results emerge from the analysis, each deserving a comprehensive treatment, are mentioned here very briefly:

1. For *departing* objects relative to the laboratory the mass-energy equivalence derived by the theory, is found to be $0.04508497 m_0 c^2$, which is exactly half Hardy's quantum coupling constant
2. The theory suggests a novel perspective of quantum phenomena, according to which the observed wave property of matter at high energies could be interpreted as a gradual transition of normal matter and normal energy to dark matter and dark energy. Such interpretation enables a long sought-after unification between Quantum Theory, and Newtonian mechanics, without leaving 95% of the Universe completely in the dark side of our knowledge.

References

1. Albrecht A. et al. Report of the Dark Energy Task Force. *Astro. Ph.*, 2006, 0609591.
2. Turner M. S. Dark matter: Theoretical perspectives. *PNAS*, 1993, v. 90, 4827–4834.
3. Riess A. G. et al. Observational evidence from supernovae for an accelerating Universe and a cosmological constant. *Astr. J.*, 1998, v. 116, 1009–1038.
4. Perlmutter S. et al. Discovery of a supernova explosion at half the age of the Universe. *Nature*, 1998, v. 391, 51–54.
5. Riess A. G. et al. The farthest known supernova: Support for an accelerating Universe and a glimpse of the epoch of deceleration. *Ap. J.*, 2001, v. 560 (1), 49–71.
6. Linder E. V. Exploring the expansion history of the Universe. *Phys. Rev. Lett.*, 2003, v. 90, 9 [4 pages].
7. Linder E. V. Probing gravitation, dark energy, and acceleration. *Phys. Rev. D*, 2004, v. 70 (2), 023511 [11 pages].
8. Sandvik H. B., Barrow J. D., Magueijo J. A Simple Cosmology with a Varying Fine Structure Constant. *Phys. Rev. Lett.*, 2002, v. 88, 031302 [4 pages].
9. Easson D. A. Framptona P. H., Smoota G. F. Entropic Accelerating Universe. *Physics Letters B*, 2011, v. 696 (3), 273–277.
10. Alcock C. EROS and MACHO combined limits on planetary-mass dark matter in the galactic halo. *The Astrophysical Journal*, 1998, v. 499, 9–12.
11. Alcock C. MACHO project limits on black hole dark matter in the 1-30 solar mass range. *The Astrophysical Journal Letters*, 2001, v. 550 (2), 169–172.
12. Steigman G., Turner M. S. Cosmological constraints on the properties of weakly interacting massive particles. *Nuclear Physics B*, 1985, v. 253, 375–386.
13. Sivertsson S., Gondolo P. The WIMP capture process for dark stars in the early Universe. *Astrophys. J.*, 2011, v. 729, 51 [11pages].
14. Aprile E. et al. Dark Matter Results from 100 Live Days of XENON100 Data. *Phys. Rev. Lett.*, 2011, v. 107, 131302 [6 pages].
15. O'Connor J. J., Roberston E. F. Christian Andreas Doppler- MacTutor History of Mathematics archive. University of St. Andrews, 1998.
16. Maulik D. Doppler Ultrasound in Obstetrics and Gynecology. Springer-Verlag, 2005.
17. Hardy L. Quantum mechanics, local realistic theories and Lorentz-invariant realistic theories. *Phys. Rev Lett.*, 1992, v. 68, 2981–2984.
18. Hardy L. Nonlocality of a single photon revisited. *Phys. Rev. Lett.*, 1994, v. 73, 2279–2283.
19. Penrose E. The Road to Reality. Jonathan Cape, London, 2004.
20. Coldea R., Tennant D. A., Wheeler E. M., Wawrzynska E., Prabhakaran D., Telling M., Habicht K., Smeibidl P., Kiefer K. Quantum criticality in an Ising chain: Experimental evidence for emergent E8 symmetry. *Science*, 2010, v. 327 (5962), 177–180.
21. Adams J. F. Lectures on exceptional Lie groups. Chicago Lectures in Mathematics, University of Chicago Press, 1996.

Appendix

1. The time transformation for the round-trip
2. Derivation of the density and energy transformations
3. The relationship between velocity (β) and redshift (z)
4. References

1 The time transformation for the round-trip

$$t = t_{Depart} + t_{Arrive} = \left(\frac{1}{1-\beta} + \frac{1}{1+\beta}\right)t' = \left(\frac{2}{1-\beta^2}\right)t', \quad (A1)$$

or,

$$\hat{t} = \frac{t}{t'} = \frac{2}{1-\beta^2}\,. \quad (A2)$$

Figure A1 depicts the relative time $\hat{t}$ as a function of β for the one-way and round trip. The dashed lines depict the corresponding predictions of *SR*.

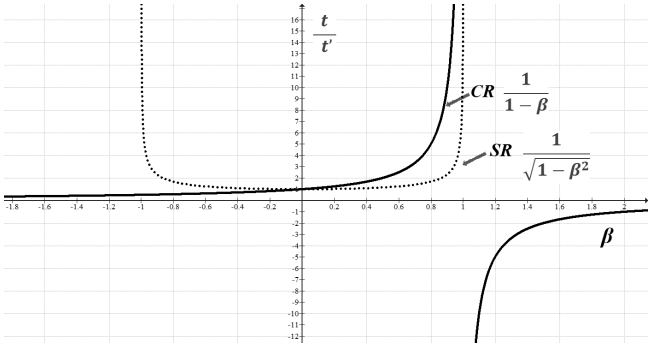

Fig. A1a

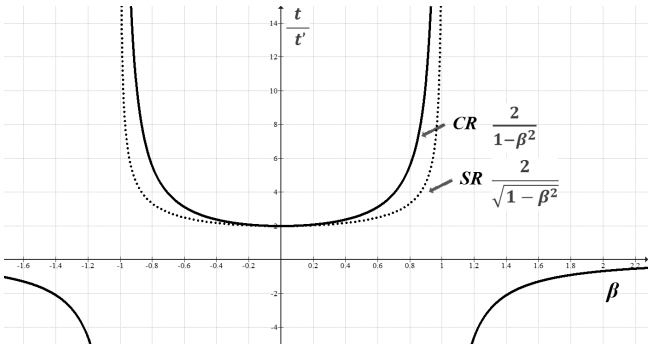

Fig. A1b

Fig. A1: Time transformations for the one-way (Fig. A1a) and round trip (Fig. A1b). The dashed lines depict the corresponding *SR* results.

2 Derivation of the density and energy transformations

To derive the density and kinetic energy transformation, consider the two frames of reference F and F' shown in Figure A2. Suppose that the two frames are moving relative to each other at a constant velocity v.

Consider a uniform cylindrical body of rest mass $m' = m_0$ and length $l' = l_0$ placed in F' along its travel direction. Suppose that at time t_1 the body leaves point x_1 (x'_1 in F') and moves with constant velocity v in the the $+x$ direction, until it reaches point x_2 (x'_2 in F') in time t_2 (t'_2 in F').

The body's density in the internal frame F' is given by:

$$\rho' = \frac{m_0}{Al_0}\,, \tag{A4}$$

where A is the area of the body's cross section, perpendicular to the direction of movement. In F the density is given by $\rho = \frac{m_0}{Al}$, where l is the object's length in F. Using the distance transformation l could be written as:

$$l = l_0\frac{1+\beta}{1-\beta}\,, \tag{A5}$$

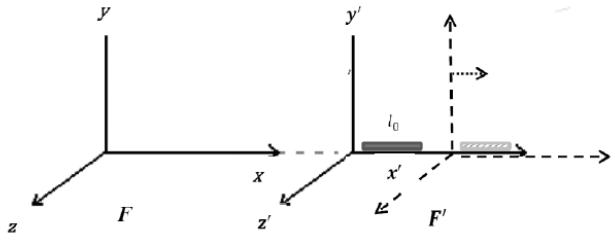

Fig. A2: Two observers in two reference frames, moving with velocity v with respect to each other

which yields:

$$\rho = \frac{m_0}{Al} = \frac{m_0}{Al_0}\frac{1-\beta}{1+\beta} = \rho'\frac{1-\beta}{1+\beta}\,,$$

or:

$$\frac{\rho}{\rho'} = \frac{1-\beta}{1+\beta}\,. \tag{A6}$$

Since the radius of the moving cylinder is perpendicular to the direction of motion, an observer at the internal frame F will measure a cylinder radius of $\Delta r = \Delta r_0$. The kinetic energy of a *unit of volume* is given by:

$$E = \frac{1}{2}\rho v^2 = \frac{1}{2}\rho_0\frac{1-\beta}{1+\beta}v^2,$$

or:

$$E = \frac{1}{2}\rho_0 c^2\beta^2\frac{1-\beta}{1+\beta}\,. \tag{A7}$$

And the energy for a departing particle of rest mass m_0 is given by:

$$E = \frac{1}{2}m_0 c^2\beta^2\frac{1-\beta}{1+\beta}\,. \tag{A8}$$

To calculate the value $\beta = \beta_{cr.}$ which satisfies $E = E_{max}$ we derive $\beta^2\frac{1-\beta}{1+\beta}$ with respect to β and equate the derivative to zero. This yields:

$$\begin{aligned}\frac{d}{d\beta}\left(\beta^2\frac{1-\beta}{1+\beta}\right) &= 2\beta\frac{1-\beta}{1+\beta} + \\ &+ \beta^2\frac{[(1+\beta)(-1)-(1-\beta)(1)]}{(1+\beta)^2} = \\ &= 2\beta\frac{(1-\beta^2-\beta)}{(1+\beta)^2} = 0\end{aligned} \tag{A9}$$

for $\beta \neq 0$ and we get:

$$\beta^2 + \beta - 1 = 0\,, \tag{A10}$$

which yields:

$$\beta_1 = -\varphi = -\frac{\sqrt{5}+1}{2} \approx -1.618 \tag{A11}$$

and

$$\beta_2 = \varphi - 1 = \frac{1}{\varphi} = \frac{\sqrt{5}-1}{2} \approx 0.618\,, \qquad (A12)$$

where φ is the Golden Ratio defined as: $\varphi = \frac{\sqrt{5}+1}{2}$ [A1-A3]. This is a striking result given the properties of this phenomenal number, due to its importance, together with the Fibonacci numbers, in mathematics, aesthetics, art, music, and more and its key role in nature, including the structure of plants, animals, the human body, human DNA [A1-A8] and brain waves [A9-A12] and in physics [A13]. The maximal kinetic energy at $\beta \approx 0.618$ is equal to:

$$E_{max} = \frac{1}{2} m_0 c^2 (\varphi - 1)^2 \frac{1-(\varphi-1)}{1+(\varphi-1)} = \\ = \frac{1}{2} m_0 c^2 (\varphi - 1)^2 \frac{2-\varphi}{\varphi}\,. \qquad (A13)$$

The term $\varphi - 1$ could be written as: $\varphi - 1 = \frac{\sqrt{5}+1}{2} - 1 = \\ = \frac{(\sqrt{5}+1)-2}{2} = \frac{\sqrt{5}-1}{2}$ Multiplying the numerator and denominator by $\frac{\sqrt{5}+1}{\sqrt{5}+1}$ yields:

$$\varphi - 1 = \frac{\sqrt{5}-1}{2}\frac{\sqrt{5}+1}{\sqrt{5}+1} = \frac{5-1}{2\sqrt{5}+1} = \\ = \frac{2}{\sqrt{5}+1} = \frac{1}{\frac{\sqrt{5}+1}{2}} = \frac{1}{\varphi}\,. \qquad (A14)$$

Eq. (A14) could be rewritten as:

$$E_{max} = \frac{1}{2} m_0 c^2 \frac{(2-\varphi)}{\varphi^3}\,. \qquad (A15)$$

Substituting $\varphi = \frac{\sqrt{5}+1}{2}$ we obtain:

$$E_{max} \approx 0.04508497 m_0 c^2. \qquad (A16)$$

3 The relationship between velocity and redshift

Redshift could be described as the relative difference between the observed and emitted wavelengths (or frequency). Let λ represents wavelength and f represents frequency ($\lambda f = c$ where c is the speed of light), then the redshift z is given by:

$$z = \frac{\lambda_r - \lambda_s}{\lambda_s} \qquad (\text{or} \quad z = \frac{f_s - f_r}{f_r})\,, \qquad (A17)$$

where $\lambda_s(f_s)$ is the wavelength (frequency) measured at the source and $\lambda_r(f_r)$ is the wavelength (frequency) measured at the receiver's laboratory.

Substituting $f_s = \frac{1}{t_s}$ and $f_r = \frac{1}{t_r}$ in (A17) above we obtain

$$z = \frac{f_s - f_r}{f_r} = \frac{\frac{1}{t_s} - \frac{1}{t_r}}{\frac{1}{t_r}} = \frac{t_r - t_s}{t_s} = \frac{t_r}{t_s} - 1. \qquad (A18)$$

But from Eq. 2 we have:

$$\frac{t_r}{t_s} = \frac{1}{1-\beta}\,. \qquad (A19)$$

Thus:

$$z = \frac{1}{1-\beta} - 1 = \frac{\beta}{1-\beta} \qquad (A20)$$

and

$$\beta = \frac{v}{c} = \frac{z}{1+z}\,. \qquad (A21)$$

4 References

A1. Olsen S. The Golden Section. New York, Walker & Co, 2006.

A2. Livio M. The Golden Ratio: The Story of Phi, the World's Most Astonishing Number. New York, Broadway Books, 2002.

A3. Huntley H. E. The Divine Proportion: a Study in Mathematical Beauty. New York, Dover, 1970.

A4. Stakhov A. Fundamentals of a new kind of mathematics based on the Golden Section. *Chaos, Solitions Fractals*, 2006, v. 27, 1124–1146.

A5. Pittard N., Ewing M., Jevons C. Aesthetic theory and logo design: examining consumer response to proportion across cultures. International Marketing Review, 2007, v. 24 (4), 457–73.

A6. Hammel G. T. Fascinating Fibonaccis: Mystery and Magic in Numbers. Dale Seymour Publications, 1987.

A7. Hammel G. T., Vaughan K. C. Math and Music: Harmonious Connections. Dale Seymour Publications, 1995.

A8. Klar A. J. S. Fibonacci's flowers. *Nature*, 2002, v. 417, 595.

A9. Weiss H., Weiss V. The golden mean as clock cycle of brain waves. *Chaos, Solitions Fractal*, 2003, v. 18 (4), 643–652.

A10. Conte E., Khrennikov A., Federici A., Zbilut J. P. Variability of brain waves: A new method based on a fractal variance function and Random Matrix Theory. *Chaos, Solitions Fractals*, 2009, v. 41 (5), 2790–2800.

A11. Weiss H., Weiss V. When frequencies never synchronize: the golden mean and the resting EEG. *Br. Res.*, 2010, v. 1335 (4), 91–102.

A12. Roopun A. K., Kramer M. A., Carracedo L. M., Kaiser M., Davies C. H., Traub R. D., Kopell N. J., Whittington M. A. Temporal interactions between cortical rhythms. *Front. Neuro.*, 2008, v. 2 (2), 145–154.

A13. Coldea R., Tennant D. A., Wheeler E. M., Wawrzynska E., Prabhakaran D., Telling M., Habicht K., Smeibidl P., Kiefer K. Quantum criticality in an Ising chain: Experimental evidence for emergent E8 symmetry. *Science*, 2010, v. 327 (5962), 177–180.

Submitted on: April 10, 2013 / Revised submission on: April 30, 2013
Accepted on May 01, 2013
After corrections: September 13, 2013

A New Model of Black Hole Formation

Gordon D. Thayer

1169 Carmella Circle, Sarasota, Florida 34243 U.S.A. E-mail: gdthayer@verizon.net

The formation of a black hole and its event horizon are described. Conclusions, which are the result of a thought experiment, show that Schwarzschild [1] was correct: A singularity develops at the event horizon of a newly-formed black hole. The intense gravitational field that forms near the event horizon results in the mass-energy of the black hole accumulating in a layer just inside the event horizon, rather than collapsing into a central singularity.

1 Introduction

This article describes the formation of a black hole and the physics of event horizon formation. In early 1916, a German physicist, Karl Schwarzschild, published a short paper in which he gave a solution to Einstein's general relativity field equations for spherically symmetric objects. Schwarzschild's solution "contains a coordinate singularity on a surface that is now named after him. In Schwarzschild coordinates, *this singularity lies on the sphere of points at a particular radius, called the Schwarzschild radius*" [1] (emphasis added). The significance of this paper has not been generally appreciated, although it led physicists eventually to accept black holes as real physical objects. Many black holes have been detected in recent years using astronomical techniques. But physicists in general have concluded that the singularity lies at the center of the black hole rather than on its event horizon. They have mostly ignored the results of Schwarzschild, who found that the singularity occurred at the event horizon itself rather than at the center of the spherical space enclosed by the event horizon. In this article I show by means of a suitably chosen thought experiment that Schwarzschild was correct.

2 A collapsing star

Following the occurrence of a Type 1a supernova, a neutron star is usually formed. For neutron stars with a mass greater than the Tolman-Oppenheimer-Volkoff limit (about 3 to 4 solar masses), the star will collapse to form a black hole. We need to follow the history of some points on and within the collapsing star in order to find out what really happens when a black hole is formed. To establish some boundary conditions, note that a point at the center of the collapsing star will not move with respect to a coordinate system centered on the star; the center of the system does not participate in the collapse. Of more interest is a point on the surface of the collapsing star. This point will have a velocity vector directed toward the center of the star with a speed that depends on the time from the initiation of collapse until the formation of the event horizon, at which time its speed is assumed to be the speed of light, c. Assume that a point halfway between the surface and the center will also have an inwardly directed velocity with half the speed of the surface point. In other words, the contraction is radially linear. Some departure from this linearity will not severely affect my conclusions.

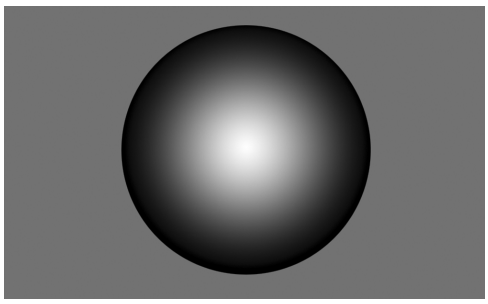

Fig. 1: Radial velocities in a collapsing star.

Figure 1 shows qualitatively what these radial velocities look like. The size of the star in the illustration is assumed to be approaching the Schwarzschild radius. The black colors indicate high radial velocity and white indicates small or zero velocities. The figure was constructed using the gradient tool in Photoshop and is linear in value from the center to the outer boundary. In reality, the darkest black should be confined to the very outer edges of the star and most of the interior should be either white or light gray. Nevertheless, the picture does give a good idea of the kind of radial velocities one would find in the cross-section of a collapsing star.

Figure 2 shows the situation at the moment when the event horizon forms. Note that the points at $0.995\,R_s$, where R_s is the Schwarzschild radius, have 10 times their normal, or rest, mass. The asymptote on the right goes to infinity at the Schwarzschild radius, $R = 1.0$ in the illustration. This is the singularity that Karl Schwarzschild discovered when he solved Einstein's field equations for a symmetrical, non-rotating body. The equation used to plot the points for the mass as a function of the radius is:

$$\frac{m}{m_0} = \frac{1}{\sqrt{1 - v^2/c^2}} \equiv \frac{1}{\sqrt{1 - R^2}}, \quad 0 \leqq R < 1. \qquad (1)$$

The validity of this special relativity equation under the conditions in the formation of an event horizon is unsure, but since a singularity is a singularity, and this equation defines one for $v = c$, it is likely as good as some other measure.

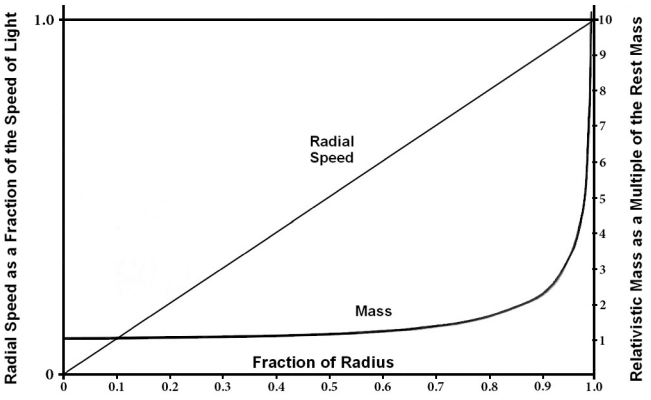

Fig. 2: Mass distribution in a newly-formed black hole. Drawing by the author.

The essential point is that most of the mass will be concentrated near the event horizon as soon as it forms. Thus the gravitational field will be quickly reversed, and with it, the velocity field inside the event horizon. Particles in the interior of the new black hole will be strongly attracted to the event horizon, since that is where most of the mass is located. This implies that the entire mass of the collapsed star could end up in a shallow region just inside the event horizon. There is no way to determine from the outside whether or not this happens.

In this scenario, the mass M is contained in a very thin layer at the radius R and the interior is empty. But how does it get there? According to Susskind [2, see p. 238] anything that impacts the event horizon of a black hole is absorbed by it, spreading over the entire extent of the event horizon the way a drop of ink dissolves rapidly in a basin of warm water. What if the event horizon itself comprises all of the mass contained in the black hole, held in a layer perhaps one Planck length in thickness? (Admittedly, that's a guess on my part.) From the outside, it would still behave like a black hole. All differences would be on the inside.

In my model the material of a collapsing star would, as soon as it has compacted enough to form a black hole, begin to migrate to the event horizon, like iron filings attracted to a magnet. The only place where the gravity of the material comprising the event horizon layer is neutral would be the exact, precise center of the black hole. But even so small a particle as a hadron would, sooner or later, wander off center — if for no other reason, because of the Heisenberg uncertainty principle. It would then be instantly attracted to the event horizon and would stick there like a bug on fly paper. Eventually the entire inside of the black hole would be empty. The layer comprising the event horizon layer may be extremely thin, but it is most definitely not a singularity, a mere mathematical point.

I recently discarded this possibility, but it appears that I may have been too hasty in doing so.

3 What happens to the matter in a black hole?

In this reconsidered theory, the singularity at the event horizon is only mathematical, not real. The mass of the collapsed star is contained in a thin layer just inside the event horizon, perhaps only a single Planck length thick. There is an external complement to this idea. Leonard Susskind [2, see pp. 233–234] writes:

> The only [solution] consistent with the laws of physics would be to assume that some kind of super-heated layer exists just above the horizon, perhaps no more than a Planck length thick... the layer must be composed of tiny objects, very likely no bigger than the Planck length. The hot layer would absorb anything that fell onto the horizon, just like drops of ink dissolving in water... This hot layer of stuff needed a name. Astrophysicists had already coined the name that I eventually settled on... They had used the idea of an imaginary membrane covering the black hole just above its horizon to analyze certain electrical properties of black holes. [They] had called this imaginary surface the stretched horizon, but I was proposing a real layer of stuff, located a Planck length above the horizon, not an imaginary surface.
>
> I liked the sound of "stretched horizon" and adopted it for my own purposes. Today the stretched horizon is a standard concept in black hole physics. It means the thin layer of hot microscopic "degrees of freedom" located about one Planck distance above the horizon.

I propose the name "Shell Theory" for my explanation of black hole formation.* This theory posits a one-to-one correspondence between the bits of entropy on the surface of the event horizon of a black hole and the particles of the collapsed star in the shell layer just inside the event horizon. The gravitational field and other external properties of the black hole will be exactly the same as if an infinite singularity existed at the center, because the amount of mass-energy in each case will be identical. All that is necessary for this condition to be true is that the distribution of mass inside the event horizon is spherically symmetrical. The shell theory has the same spherical symmetry as conventional theory with a singularity at the center of the black hole.

In the shell theory evolution of a black hole, the collapsing of the remnant star must stop as soon as the event horizon is formed. The reversal would start at a time somewhat prior to the formation of the event horizon. In figure 1 it is apparent that even before the outer layer of particles achieves a velocity magnitude equal to the speed of light, the distribution of

*For the purposes of this article, a "shell" is defined as the volume enclosed between concentric spheres of different radii.

mass within the collapsing object would favor the outer layers over the inner layers. This differential in the gravitational field would build up rapidly as the size of the collapsing star approached the Schwarzschild limit, so it would not be an instantaneous reversal.

The mass of a differential shell from the collapsing star as a function of the radius, assuming that the radial velocity of a point inside the object is a linear function of the radius up to the limit of $v = c$, at $R = R_s$, is:

$$d\frac{m}{m_0} = 4\pi R^2 \frac{m}{m_0}\, dR\,, \tag{2}$$

where

$$\frac{m}{m_0} = \frac{1}{\sqrt{1 - R^2}}\,. \tag{3}$$

Therefore the total relative mass of a spherical shell is given by the integral:

$$\frac{m}{m_0} = \int \frac{4\pi R^2 dR}{\sqrt{1 - R^2}} = 4\pi \left[\frac{1}{2}\sin^{-1}(R) - \frac{1}{2} R\sqrt{1 - R^2}\right]. \tag{4}$$

This result must be evaluated at three points: $R = 0$; $R = R$; and $R = R_s$. The result for $R = 0$ is simple: 0. For $R = R_s$ the term $(1 - R^2)$ becomes zero, and $\sin^{-1}(1)$ is $\frac{\pi}{2}$; so the result for $R = R_s$ is $\frac{\pi}{4}\,(\times 4\pi)$. Subtracting the two solutions from each other (ignoring the common factor of 4π) and setting the results equal to each other — so that we obtain the radius within which and without which there is equal mass — we have, after rearranging terms, the equation:

$$\sin^{-1}(R) = \frac{\pi}{4} + R\sqrt{1 - R^2}. \tag{5}$$

This equation, (5), is difficult to evaluate in closed form, but the result can be obtained easily through the process of successive iterations. The solution is approximately $R = 0.915$ (the difference between the two sides of the equation is 9×10^{-4} out of 1.155), meaning that the outer 8.5% of the sphere contains as much relativistic mass as the entire inner 91.5%.* This amply demonstrates that what was initially the inward implosion of a neutron star will now be a radially outward "explosion" within the confines of the event horizon — the surface implied by Schwarzschild's results.

4 Results and discussion

The likely end result will be that all of the mass-energy of a collapsed star ends up confined to a very thin layer — probably only one Planck length thick — just inside the event horizon. There may be a " black hole" there, but its matter will not be located in an infinitely dense singularity at the center point.

Also notice that for a solid body of uniform density, the gravitational field outside the surface is inversely proportional to the square of the distance from the center of the body, but for points inside the body the gravitational field is linear, diminishing to zero at the center. This reinforces the assumption that the collapse of the neutron star should be linear in nature. The effect as the radius of the shrinking star approaches and attains the Schwarzschild radius is to change this linear gravitational potential into a hyperbolic gravitational field, asymptotic to infinity at R_s.

The singularity at the Schwarzschild radius, or event horizon, is mathematical only and does not affect any real particles. The event horizon is described by a metric of points distributed over a spherical manifold, and the term "point mass" is an oxymoron since a point cannot have mass or any other physical property. It is nothing more than a mathematical position in space-time. In this context, note that the integration in equation (4) does not diverge at $R = R_s$, as it would if there were a true infinity at that point.

Where I have written the word "point" or "points", this term should not be taken literally. The reader should imagine a tiny amount of matter, perhaps a cubic Planck length (Planck volume) in size, located at a particular point in space-time. An actual point has no dimensions and therefore cannot have mass or any other physical property. The Planck volume is believed by many to be a quantum unit of space.

Submitted on July 8, 2013 / Accepted on July 14, 2013

References

1. Schwarzschild K. On the gravitational field of a sphere of incompressible fluid according to Einstein's theory. Communicated February 24th, 1916. *Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften zu Berlin, Phys.-Math. Klasse 1916*, 424–434. Translated from the German in 2008 by Larissa Borissova and Dmitri Rabounski. *The Abraham Zelmanov Journal*, 2008, volume 1, pages 20–32
2. Susskind L. The Black Hole War. New York, Little, Brown and Company, 2008.

*A more precise result is $0.914554 \pm 2 \times 10^{-6}$.

Dynamical 3-Space: Gravitational Wave Detection and the Shnoll Effect

David P. Rothall* and Reginald T. Cahill‡

School of Chemical and Physical Sciences, Flinders University, Adelaide 5001, Australia.
E-mail: *David.Rothall@flinders.edu.au, ‡Reg.Cahill@flinders.edu.au

Shnoll has investigated the non-Poisson scatter of rate measurements in various phenomena such as biological and chemical reactions, radioactive decay, photodiode current leakage and germanium semiconductor noise, and attributed the scatter to cosmophysical factors. While Shnoll didn't pinpoint the nature of the cosmophysical factors the Process Physics model of reality leads to a description of space, which is dynamic and fractal and exhibits reverberation effects, and which offers an explanation for the scattering anomaly. The work presented here shows a new way of generating the effects Shnoll discovered, through studying the phase difference of RF EM waves travelling through a dual coaxial cable Gravitational Wave Detector experiment.

1 Introduction – Shnoll effect

Over sixty years ago Simon Shnoll discovered a scatter anomaly in the measurements of the reaction rates of ATP-ase in actomyosin solutions over time that could not be explained [1]. Extensive research into this scatter anomaly lead to the conclusion that the reaction rates of the protein solution not only varied with time, but followed a distribution with preferred (discrete) values instead of a typical Poisson distribution. Over the following decades it was found that quite different phenomena also displayed similar scatter anomalies, ranging from chemical reactions to α-radiation activity in ^{239}Pu decay, photomultiplier dark noise and semiconductor noise fluctuations [2]. Shnoll's investigation of the scatter anomaly (referred to here as the Shnoll effect), between May 28 - June 01, 2004, produced 352,980 successive measurements of the α decay of a ^{239}Pu source [1]. Radioactive decay is considered to be a stochastic process, i.e. a random process with no preferred frequencies, and hence follows Poisson statistics. Fig. 1 is a layer histogram taken from Shnoll's data, with layer lines taken every 6000 measurements. The y-axis represents the frequency of decay rates and the x-axis is the number of decays per second - the decay rate. Over time the layer lines of the histogram exhibit a fine structure which become more prominent with more measurements, instead of canceling out as in the case of a typical Poisson distribution. This suggests that the radioactivity of ^{239}Pu takes on discrete values, and is not completely random.

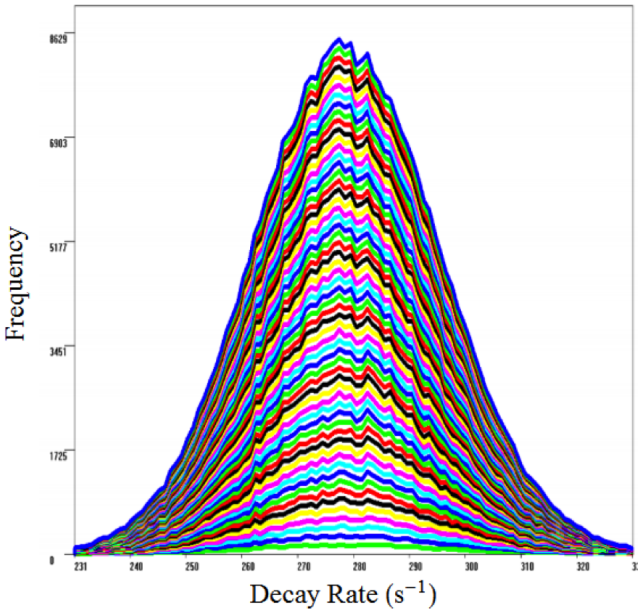

Fig. 1: Non-Poisson distribution of 352,980 measurements of ^{239}Pu α decay by Shnoll performed in 2004 (Fig. 2-2 of [1]). The layered histograms are taken every 6000 measurements. The x-axis denotes the number of decay events per second and the y-axis is the frequency of decay rate measurements.

Upon further study it was found that not only did the distribution (histogram) shapes vary over time, but the histogram shapes also correlated between different experiments run in parallel, regardless of whether they were located in the same laboratory or separated by thousands of kilometres. This was referred to as absolute time synchronism. Local-time synchronism was also observed, where histogram shapes of one experiment matched those of another with a time delay corresponding to the difference in longitudes of the two locations of the experiments (i.e. as the Earth rotates). A "near zone" effect was also discovered, where consecutive histograms in time of an individual experiment were found to be most similar in shape, regardless of the time scale used to generate the histograms, indicating the fractal nature of the scattering anomaly. The main conclusions drawn from Shnoll's research was that the consistency of the "scattering of results" of measurements in a time series arise due to inhomogeneities in the "space-time continuum" [1, 7]. These inhomogeneities are "caused by the movement of an object in the inhomogeneous gravitational field", e.g. as the Earth rotates/orbits the Sun, as the moon orbits the Earth, etc. While these inhomogeneities were not characterized by Shnoll there is a remark-

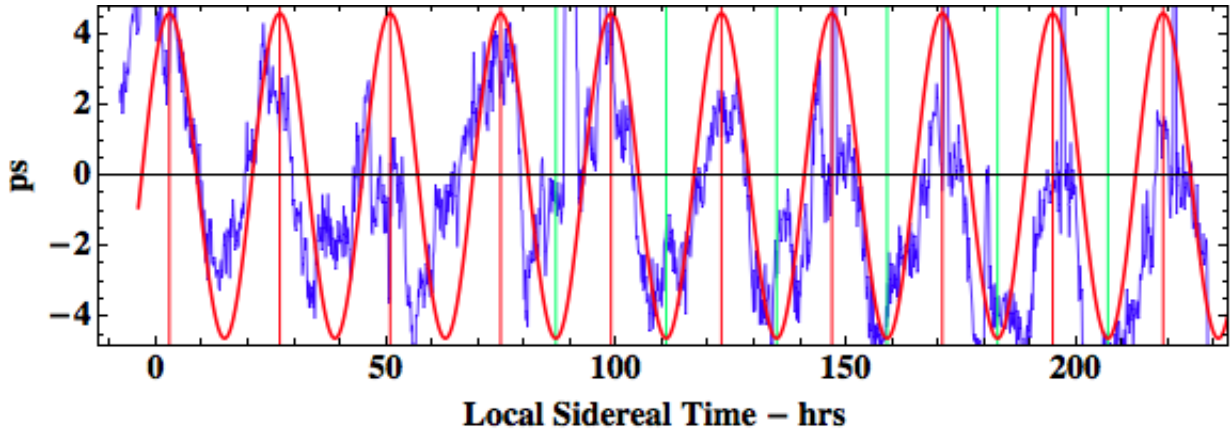

Fig. 2: Reproduction of Fig. 8 (top) from [5] showing the travel time differences (ps) between the two coaxial cable circuits in [5] plotted against local sidereal time, for the duration March 4 - 12, 2012. The smooth sine wave is a prediction made from the Dynamical 3-Space theory using NASA spacecraft Earth-flyby Doppler shift data.

able amount of evidence supporting this conclusion.

2 Dynamical 3-Space

An alternative explanation of the Shnoll effect has been proposed using an alternative theory known as dynamical 3-space theory; see Process Physics [3]. This arose from modeling time as a non-geometric process, i.e. keeping space and time as separate phenomena, and leads to a description of space which is itself dynamic and fractal in nature. It uses a uniquely determined generalisation of Newtonian Gravity expressed in terms of a velocity field $\boldsymbol{v}(\boldsymbol{r}, t)$, defined relative to an observer at space label coordinate $\boldsymbol{r}$, rather than the original gravitational acceleration field. The dynamics of space in the absence of vorticity, $\nabla \times \boldsymbol{v} = \boldsymbol{0}$, becomes*

$$\nabla\cdot\left(\frac{\partial \boldsymbol{v}}{\partial t} + (\boldsymbol{v}\cdot\nabla)\boldsymbol{v}\right) + \frac{5\alpha}{4}\left((trD)^2 - tr(D^2)\right) = -4\pi G\rho \qquad (1)$$

$$D_{ij} = \frac{\partial v_i}{\partial x_j} \qquad (2)$$

where $\rho(\boldsymbol{r}, t)$ is the usual matter density. The 1st term involves the Euler constituent acceleration, while the α-term describes a significant self interaction of space. Laboratory, geophysical and astronomical data suggest that α is the fine structure constant $\approx 1/137$. This velocity field corresponds to a space flow which has been detected in all of the experiments listed in section 3. In the spherically symmetric case and in the absence of matter $\rho = 0$, (2) contains solutions for black holes (spatial inflows) and an expanding universe (Hubble expansion) along with that for black holes embedded in an expanding universe [4]. (2) also contains solutions for the inflow of space into a matter density. Perturbing the spatial inflow into matter (i.e. simulating gravitational waves) was shown recently to produce reverberations in which the wave generates trailing copies of itself [8]. This reverberation effect is caused by the non-linear nature of the flow dynamics evident in (2) and will be shown in the coaxial cable data discussed in section 3.

*The α term in (2) has recently been changed due to a numerical error found in the analysis of borehole data. All solutions are also altered by these factors. (2) also contains higher order derivative terms - see [4] .

3 2012 Dual RF coaxial cable experiment

The Dynamical 3-Space theory was applied to an experiment which studied the radio frequency (RF) electromagnetic (EM) speed anisotropy, or ultimately the absolute motion of Earth through space. The effect of absolute motion has previously been studied using the results from Michelson - Morley, Miller, and DeWitte experiments [5]. These results are in remarkable agreement with the velocity of absolute motion of the Earth determined from NASA spacecraft Earth-flyby Doppler shift data all revealing a light/EM speed anisotropy of some 486 km/s in the direction RA=4.3^h, Dec = −75.0° [6]. The actual daily average velocity varies with days of the year because of the orbital motion of the Earth - the aberration effect discovered by Miller, but also shows fluctuations over time. The dual RF coaxial cable experiment, performed from March 4 - 12, 2012, measures the travel time difference of two RF signals propagating through dual coaxial cables [5]. The key effect in this 1st order in v/c experiment is the absence of the Fresnel drag effect in RF coaxial cables at a sufficiently low frequency. The experiment is designed such that one RF signal travels through one type of coaxial cable and returns via another type of cable, while the other signal does exactly the opposite. The cables are bound together such that any travel time effects due to temperature changes cancel as both cables are affected equally. Fig. 2 is a reproduction of the data obtained from the experiment in March 2012 where the travel time difference between the RF signals is plotted against sidereal time. The data is fitted, smooth curve, using

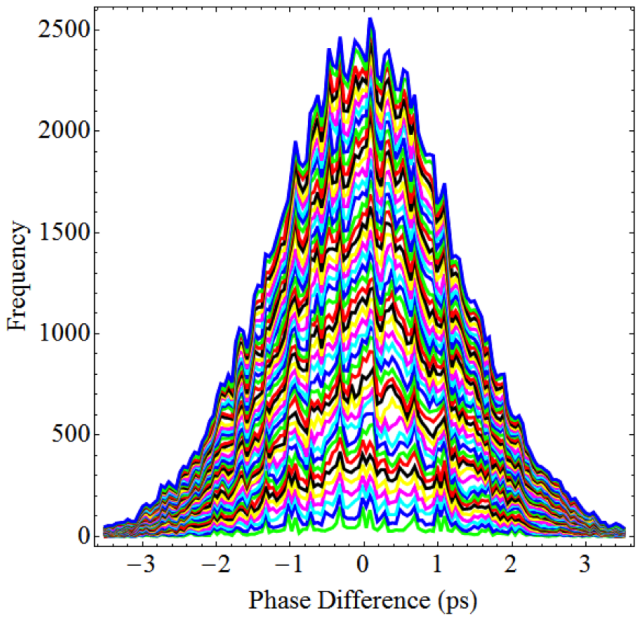

Fig. 3: Non-Poisson distribution of 155,520 measurements of the travel time difference (ps) observed between the two coaxial cable circuits of [5] from Mar 04, 2012 to Mar 12, 2012 in Adelaide. The layered histograms are taken every 3350 measurements to show a comparison with that of Fig. 1.

predictions from the NASA spacecraft Earth-flyby Doppler shift data, where a flow of space traveling at a speed of 499 km/s and direction RA=2.75^h, Dec = −77° predicts the overlaid sine wave, with dynamic range ∼8 ps. The Earth rotation effect, with respect to the galaxy, can be observed from the data, as well as turbulence effects. Turbulence effects are beginning to be characterized, and can be shown to correspond to what are, conventionally known as gravitational waves, although not those implied by General Relativity, but more precisely are revealing a fractal structure to dynamical 3-space, as illustrated in Fig. 4.

A Fast Fourier Transform of the data in Fig. 2 was taken to remove the Earth rotation effect (i.e. low frequency effects), and then a histogram taken of the resultant 155,520 measurements (after inverse FFT) to generate the layered histogram plot shown in Fig. 3. Layer lines are inserted every 3350 measurements to show a comparison with the Shnoll plot in Fig. 1. Fig. 3 is remarkably comparable to Fig. 1, thus suggesting that the Shnoll effect is also present in the coaxial cable EM anisotropy experiments. The structure observed appears to build up over time instead of cancelling out. It appears slightly noisier but this may be due to the fewer data points obtained than Shnoll (352,980 measurements). The structure observed is found to persist regardless of the time scale used for the phase difference, suggesting that the phenomenon causing this has a fractal nature as depicted in Fig.4. If this is indeed caused by a dynamical and fractal 3-space

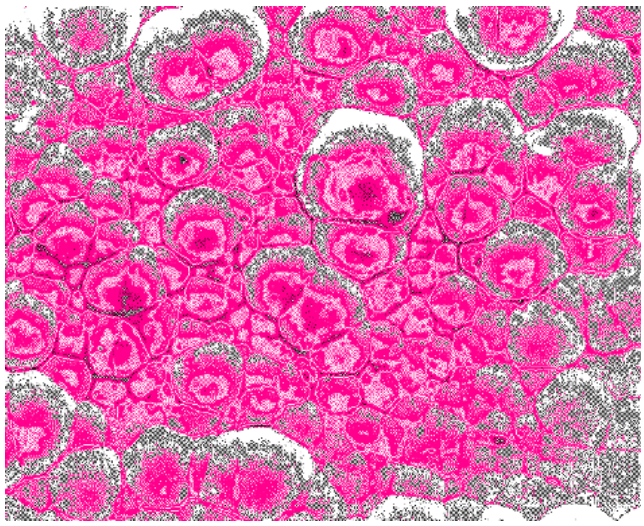

Fig. 4: Representation of the fractal wave data as revealing the fractal textured structure of the 3-space, with cells of space having slightly different velocities and continually changing, and moving with respect to the earth with a speed of ∼500 km/s.

then the persisting structure observed in Figures 1 and 3 correspond to regions of space passing the Earth that have preferred/discrete velocities, and not random ones, as randomly distributed velocities would result in a Poisson distribution, i.e. no features. A likely explanation for this is that the gravitational waves propagating in the 3-space inflow of the Earth or Sun could become phase locked due to the relative locations of massive objects. This would cause reverberation effects, i.e. regions of space which have the same speed and direction, which then repeat over time. The reverberations would be detectable in many experiments such as EM anisotropy, radiation decay, semiconductor noise generation, etc. and could in the future be used to further characterize the dynamics of space.

4 Conclusion

The data from a dual RF coaxial-cable / EM anisotropy - gravitatonal wave experiment displays the effect Shnoll observed previously in radioactivity experiments. It is suggested that these two experiments (along with other work by Shnoll) are caused by the fractal nature of space, together with the reverberation effect from gravitational waves, as predicted by the Dynamical 3-Space theory.

5 Acknowledgments

Special thanks to Professor Simon Shnoll for permission to use data from his work - see [1] for details. This report is from the Flinders University Gravitational Wave Detector Project, funded by an Australian Research Council Discovery Grant: *Development and Study of a New Theory of Gravity*.

Submitted on July 9, 2013 / Accepted on July 10, 2013

References

1. Shnoll S. E. Cosmophysical Factors in Stochastic Processes. American Research Press, Rehoboth, New Mexico, USA, 2012. Available from Progress in Physics website.
2. Shnoll S. E. Experimental investigations of local-time effect existence on laboratory scale and heterogeneity of space-time. *Progress in Physics*, 2007, v. 1, 64–69.
3. Cahill R. T. Process Physics: From Information Theory to Quantum Space and Matter. Nova Science Pub., New York, 2005.
4. Rothall D. P. and Cahill R. T. Dynamical 3-Space: Black Holes in an Expanding Universe. *Progress in Physics*, 2013, v. 4, 25–31.
5. Cahill R. T. Characterisation of Low Frequency Gravitational Waves from Dual RF Coaxial-Cable Detector: Fractal Textured Dynamical 3-Space. *Progress in Physics*, 2012, v. 3, 3–10.
6. Cahill R. T. Combining NASA/JPL One-Way Optical-Fiber Light-Speed Data with Spacecraft Earth-Flyby Doppler-Shift Data to Characterise 3-Space Flow. *Progress in Physics*, 2009, v. 4, 50–64.
7. Shnoll, S. E., Zenchenko T. A., Zenchenko K. I., Pozharskii E. V., Kolombet V. A. and Konradov A. A. Regular variation of the fine structure of statistical distributions as a consequence of cosmophysical agents. *Phys. Usp.*, 2000, v. 43 (2), 205–209.
8. Cahill R. T. and Deane S. T. Dynamical 3-Space Gravitational Waves: Reverberation Effects. *Progress in Physics*, 2013, v. 2, 9–11.

Potentialities of Revised Quantum Electrodynamics

Bo Lehnert

Alfvén Laboratory, Royal Institute of Technology, 10044 Stockholm, Sweden. E-mail: Bo.Lehnert@ee.kth.se

The potentialities of a revised quantum electrodynamic theory (RQED) earlier established by the author are reconsidered, also in respect to other fundamental theories such as those by Dirac and Higgs. The RQED theory is characterized by intrinsic linear symmetry breaking due to a nonzero divergence of the electric field strength in the vacuum state, as supported by the Zero Point Energy and the experimentally confirmed Casimir force. It includes the results of electron spin and antimatter by Dirac, as well as the rest mass of elementary particles predicted by Higgs in terms of spontaneous nonlinear symmetry breaking. It will here be put into doubt whether the approach by Higgs is the only theory which becomes necessary for explaining the particle rest masses. In addition, RQED theory leads to new results beyond those being available from the theories by Dirac, Higgs and the Standard Model, such as in applications to leptons and the photon.

1 Introduction and background

The vacuum state is not merely that of an empty space. Its energy has a nonzero ground level, the Zero Point Energy, being derived from the quantum mechanical energy states given e.g. by Schiff [1]. An example on the related vacuum fluctuations was provided by Casimir [2] who predicted that two closely spaced metal plates will attract each other. This is due to the fact that only small wavelengths can exist in the spacing, whereas the full spectrum of fluctuations exerts a net force on the outsides of the plates. The Casimir force was first demonstrated experimentally by Lamoreaux [3]. It implies that the vacuum fluctuations generate a real physical pressure and pressure gradient. Part of the quantum fluctuations also carry electric charges, as pointed out e.g. by Abbot [4]. The observed electron-positron pair formation from an energetic photon further indicates that electric charges can be created out of an electrically neutral state.

These established facts form the starting point of a revised quantum electrodynamic (RQED) theory by the author [5]. The theory is thus based on the hypothesis of a nonzero electric field divergence, $\operatorname{div}\mathbf{E} \neq 0$, in the vacuum. At the same time there is still a vanishing magnetic field divergence, $\operatorname{div}\mathbf{B} = 0$, due to the experimental fact that no magnetic monopoles have so far been observed. A nonzero electric field divergence has the following fundamental consequences [5]:

- The symmetry between the electric and magnetic fields **E** and **B** is *broken*.
- The nonzero electric charge density of a configuration with internal structure can both lead to a *net* integrated charge, and to *intrinsic charges* of both polarities.
- There exist *steady* electromagnetic states in the vacuum for which the energy density of the electromagnetic field gives rise to nonzero *rest masses* of corresponding particle models.

In the following treatise the basic field equations of RQED theory are first shortly described in Section 2. This is followed in Section 3 by a comparison to the related theories by Dirac as summarized by Morse and Feshbach [6], and that by Higgs [7]. The features and potentialities of RQED theory have earlier been described by the author [5, 8]. In Section 4 some complementary points will be presented, with special emphasis on results obtained beyond the Standard Model and not being deducible from other theories.

2 Basic field equations of Revised Quantum Electrodynamics

In four-dimensional representation the electromagnetic field equations have the general form

$$\left(\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \nabla^2\right)A_\mu = \mu_0 J_\mu \qquad \mu = 1,2,3,4 \tag{1}$$

with the four-potentials $A_\mu = (\mathbf{A}, i\phi/c)$, **A** and ϕ as the magnetic vector potential and the electrostatic potential in three-space, and the four-current

$$J_\mu = (\mathbf{j}, ic\bar{\rho}) \tag{2}$$

with **j** and $\bar{\rho}$ as electric current density and electric charge density in three-space. The form (1) is obtained from the original set of equations through a gauge transformation in which the Lorentz condition

$$\operatorname{div}\mathbf{A} + \frac{1}{c^2}\frac{\partial\phi}{\partial t} = 0 \tag{3}$$

is imposed.

The source term due to the four-current (2) in the right-hand member of (1) has to satisfy the Lorentz invariance. This implies that

$$\mathbf{j}^2 - c^2\bar{\rho}^2 = const = 0 \tag{4}$$

when J_μ is required to vanish with the charge density $\bar{\rho}$. This finally results in a four-current

$$J_\mu = \bar{\rho}\,(\mathbf{C}, ic) \tag{5}$$

where

$$\mathbf{C}^2 = c^2 \tag{6}$$

and $\mathbf{C}$ is a velocity vector with a modulus equal to the velocity constant c of light. Concerning (6) two points should be observed [5]:

- The vector $\mathbf{C}$ both includes the case of a plane wave propagating at the scalar velocity c, and three-dimensional cases such as those of a cylindrical wave where $\mathbf{C}$ has at least two spatial components. In this way (6) can be considered as an *extension* of the Lorentz invariance to three dimensions.
- Equation (6) is quadratic and leads to two solutions. These represent the two resulting *spin* directions.

In a three-dimensional representation the field equations in the vacuum now become

$$\frac{\operatorname{curl}\mathbf{B}}{\mu_0} = \varepsilon_0\,(\operatorname{div}\mathbf{E})\,\mathbf{C} + \varepsilon_0\frac{\partial\mathbf{E}}{\partial t} \tag{7}$$

$$\operatorname{curl}\mathbf{E} = -\frac{\partial\mathbf{B}}{\partial t} \tag{8}$$

where

$$\mathbf{B} = \operatorname{curl}\mathbf{A} \qquad \operatorname{div}\mathbf{B} = 0 \tag{9}$$

$$\mathbf{E} = -\nabla\phi - \frac{\partial\mathbf{A}}{\partial t} \qquad \operatorname{div}\mathbf{E} = \frac{\bar{\rho}}{\varepsilon}. \tag{10}$$

These equations are gauge invariant, as in all cases where Maxwell's equations also include source terms.

The basic features of the RQED field equations are thus specified and summarized by the following points:

- The abolished symmetry between the electric and magnetic fields leads to equations having the character of *intrinsic linear* symmetry *breaking*.
- The equations are both Lorentz and gauge *invariant*.
- There is a *source* given by the "space-charge current density" of the first term in the right-hand member of (7). Through the nonzero electric field divergence this form introduces an additional degree of freedom, leading to new physical phenomena.
- Electromagnetic *steady* states with corresponding nonzero rest masses occur on account of (7).
- *New* and *modified* wave modes arise from the extended form (6) of Lorentz invariance.
- There is full symmetry between the solutions of positive and negative polarity, thereby realizing particle models for matter as well as for *antimatter*.

As described by Schiff [1] among others, Maxwell's equations are used as a guideline for proper interpretation of conventional quantum electrodynamical theory. Thereby Heitler [9] has shown that the quantized electrodynamic equations become identical with the original classical equations in which the electromagnetic potentials and currents merely become replaced by their quantum mechanical *expectation* values. In an analogous way, this also applies to the present RQED theory.

2.1 Steady electromagnetic states

As an example on steady electromagnetic states, a particle-shaped axisymmetric configuration is now considered in a spherical frame (r, θ, φ) with a current density $j = (0, 0, C\bar{\rho})$ and a magnetic vector potential $\mathbf{A} = (0, 0, A)$. Here $C = \pm c$ represents the two spin directions. From equations (7)–(10) with $\partial/\partial t = 0$, $\partial/\partial\varphi = 0$, $\rho = r/r_0$ and r_0 standing for a characteristic radial dimension, the result becomes [5]

$$CA = -\left(\sin^2\theta\right)DF \tag{11}$$

$$\phi = -\left[1 + \left(\sin^2\theta\right)D\right]F \tag{12}$$

$$\bar{\rho} = -\frac{\varepsilon_0}{r_0^2\rho^2}D\left[1 + \left(\sin^2\theta\right)D\right]F \tag{13}$$

where

$$\begin{gathered} D = D_\rho + D_\theta \\ D_\rho = -\frac{\partial}{\partial\rho}\left(\rho^2\frac{\partial}{\partial\rho}\right) \quad D_\theta = -\frac{\partial^2}{\partial\theta^2} - \frac{\cos\theta}{\sin\theta}\frac{\partial}{\partial\theta} \end{gathered} \tag{14}$$

and there is a separable generating function

$$\begin{gathered} F\,(r,\theta) = CA - \phi = G_0 \cdot G\,(\rho,\theta) \\ G\,(\rho,\theta) = R\,(\rho)\cdot T\,(\theta)\,. \end{gathered} \tag{15}$$

With equations (11)–(15) the net electric charge q_0, magnetic moment M_0, rest mass m_0, and integrated spin s_0 are then given by

$$q_0 = 2\pi\varepsilon_0 r_0 G_0 J_q \tag{16}$$

$$M_0 = \pi\varepsilon_0 C r_0^2 G_0 J_M \tag{17}$$

$$m_0 = \frac{\pi\varepsilon_0}{c^2} r_0 G_0^2 J_m \tag{18}$$

$$s_0 = \frac{\pi\varepsilon_0 C}{c^2} r_0^2 G_0^2 J_s \tag{19}$$

where

$$J_k = \int_{\rho_k}^{\infty}\int_0^{\pi} I_k\, d\rho\, d\theta \qquad k = q, M, m, s \tag{20}$$

and I_k are differential expressions given in terms of the quantities and operators of equations (11)–(15). In the integrals (20) the radii $\rho_k = 0$ when G is convergent at $\rho = 0$, and $\rho_k \neq 0$

are small radii of circles centered at $\rho = 0$ when G is divergent at $\rho = 0$ and a special renormalisation procedure has to be applied.

The form (15) of generating function has four alternatives. When $R(\rho)$ is divergent at $\rho = 0$ and $T(\theta)$ has top-bottom symmetry, there is a nonzero net charge q_0 and magnetic moment M_0, leading to models of charged leptons. In the remaining three cases both q_0 and M_0 vanish, thereby leading to neutral leptons such as massive neutrinos.

In addition to the quantization leading to expectation values of the field vectors, relevant second quantization conditions have to be imposed on the forms (16)–(19). These concern the spin, the magnetic moment, and the total magnetic flux [5].

2.2 New and modified wave modes

Due to experimental evidence, a model representing the wave packet of an individual photon in the vacuum has to satisfy the following general requirements:

- It should have a preserved and spatially limited geometrical shape of a wave packet propagating in an undamped way and in a defined direction, even at cosmical distances.
- To limit its geometrical shape, no artificial boundaries are to be imposed on the solutions of the field equations.
- The angular momentum in the direction of propagation, the spin, should be nonzero and have the constant value $h/2\pi$.

The field equations (7)–(10) have solutions satisfying these requirements. This applies e.g. to cylindrical waves in a frame (r, φ, z) with z along the direction of propagation. For these waves the velocity vector has the form

$$\mathbf{C} = c\,(0, \cos\alpha, \sin\alpha) \tag{21}$$

with a constant angle α. Normal modes varying as $f(r)\exp[i[-\omega t + kz)]$ in an axisymmetric case lead to the dispersion relation

$$\omega = kv \qquad v = c\,(\sin\alpha) \tag{22}$$

having phase and group velocities equal to v. Expressions for the components of $\mathbf{E}$ and $\mathbf{B}$ are then obtained from the separable generating function. A wave packet of narrow line width at a main wavelength λ_0 is further formed from a spectrum of these elementary modes. This finally leads to spatially integrated quantities such as net electric charge q, magnetic moment M, total mass m, and total spin s_z. The result is as follows:

- Both q and M vanish.
- There is a finite nonzero spin

$$\mathbf{s} = \mathbf{r} \times \frac{\mathbf{S}}{c^2} \qquad \mathbf{S} = \mathbf{E} \times \mathbf{B}/\mu_0 \tag{23}$$

  where $\mathbf{r}$ is the radius vector, $\mathbf{S}$ the Poynting vector, and $s_z = h/2\pi$ for the component of $\mathbf{s}$ in the z direction.
- A finite mass

$$m = m_0/\,(\cos\alpha) \tag{24}$$

  is obtained where m_0 stands for a nonzero but very small rest mass.

This solution leads to a characteristic radial dimension $\hat{r}$ for two modes given by

$$\hat{r} = \frac{\lambda_0}{2\pi\,(\cos\alpha)} \begin{cases} 1 & \text{(25a)} \\ \varepsilon & \text{(25b)} \end{cases}$$

where (25a) refers to a convergent generating function, and (25b) to a generating function which is divergent at $r = 0$ and where a special renormalisation procedure has to be applied.

The phase and group velocities of (22) are smaller than the velocity constant c. Still this difference from c can become small enough to be *hardly distinguishable*. An example can be given by $\sin\alpha = 1 - \delta$, $0 < \delta \ll 1$, $\varepsilon = \cos\alpha$, $0 < \varepsilon \ll 1$, and $\lambda_0 = 3 \times 10^{-7}$ m for a main wavelength in the visible range. When $\delta = 10^{-10}$ this yields characteristic radii of about 3×10^{-3} m and 5×10^{-7} m due to equations (25a) and (25b).

3 Relations to other fundamental theories

It has further to be established how the present RQED approach is related to such fundamental theories as that by Dirac [6] and by Higgs [7] with the associated Standard Model of elementary particles.

3.1 The theory by Dirac

To bring wave mechanical theory into harmony with the theory of relativity, Dirac adopted a new wave equation. Then it need not to be *assumed* that the electron is spinning or turning on its axis. According to the theory the electron will have an internal angular momentum (spin), and an associated magnetic moment. In fact there are four wave functions and corresponding matrices instead of one. These alternatives thus correspond to two spin directions, and to the two possibilities of matter and antimatter, such as in the form of the electron and the positron.

As seen from the previous sections, the present RQED theory is in full correspondence with that by Dirac, in including the two spin directions as well as particles and antiparticles. But the net elementary charge, e, and the finite electron rest mass, m_e, are only included as given and assumed parameters in the theory by Dirac, whereas these quantities are *deduced* from the field equations of RQED. The latter theory also leads to other new results beyond those being available from that by Dirac.

3.2 The theory by Higgs

The Standard Model of the theory on elementary particles is based on the source-free solutions of the field equations in

the vacuum as an empty space, i.e. (1) with a vanishing right-hand member. This leads to the Hertz equations having a vanishing electric field divergence, and it results in massless particles, in contradiction with their experimentally confirmed massive counterparts.

To resolve this contradiction, Higgs [7] proposed a *spontaneous nonlinear* mechanism of symmetry *breaking* by which an unstable boson of unspecified but large nonzero rest mass is formed, having vanishing spin and electric charge. The Higgs boson then decays into a whole succession of massive elementary particles.

During many years attempts have been made to find the Higgs boson. Finally the highly advanced and imposing experiments performed by the projects ATLAS [10] and CMS [11] at CERN have bebouched into the important confirmation of an existing unstable Higgs-like boson. The latter has been found to be characterized by vanishing electric charge and spin, combined with a rest mass of about 125 GeV. It was also observed to decay rapidly into successions of particles with smaller nonzero rest masses.

However, it could here be put into doubt whether this important experimental result provides a unique confirmation of the theory by Higgs, or if the theory described in Section 2 of this paper could as well explain the results without reference to the theory by Higgs. This question can be divided into two parts, i.e. the formation of a Higgs-like particle, and its decay. The first part thus concerns formation of a particle of mass in the range of 125 GeV, having vanishing charge and spin. Equations (11)–(15) imply that massive particles can be created already from the *beginning* by the intrinsic linear broken symmetry mechanism of RQED. Among the obtained solutions there is one which is expected to become unstable, having an unspecified but nonzero and large rest mass, as well as vanishing charge and spin [12]. Such a particle of mass 125 GeV can thus be predicted. Concerning the second part of the raised question, the resulting particle would, as in all earlier known cases, decay into several other massive particles in a way being independent of and not being unique for the Higgs mechanism. In this connection it might at a first sight be argued that the Higgs-like particle obtained from RQED is not identical with that considered by Higgs. This would, however, lead to the unlikely situation of two particles having the same basic and initial data of mass, charge and spin and resulting into the same decay processes, but still not being identical.

There may finally exist a certain similarity between the source of the Higgs field and that of the Zero Point Energy of RQED.

4 New results beyond other approaches

There are results from RQED which are not deducible from the Standard Model and other fundamental theories, as being demonstrated here by a number of examples.

4.1 Models of leptons

The field equations (7)–(10) in a steady state $\partial/\partial t = 0$ lead to new results and solutions:

- Charged lepton models arise from a divergent generating function and result in a point-charge-like geometry of *small* radial dimensions, such as that of the electron.
- A *deduced* elementary electric net charge is obtained. It is located within a *narrow* parameter channel situated around the experimental value, e, and having a width of only a few percent of e.
- Through a revised renormalisation process all relevant quantum conditions and all experimental values of charge, magnetic moment, rest mass, and spin can be reproduced by the choice of *only two* free scalar parameters, the so called counter-factors.
- The magnetic field contribution to equations (7)–(10) *prevents* charged leptons from "exploding" under the action of their electrostatic eigenforce.
- There are intrinsic electric charges of both polarities in leptons, each being about an order of magnitude larger than the net elementary charge e. It results in a Coulomb interaction force between these particles, being about two orders of magnitude larger than that due to the net charge. If these conditions would also hold for quarks, the total Coulomb force would become *comparable* and *similar* to the short-range interaction of the strong force [13]. This raises the question whether the intrinsic charge force will interfere with the strong force, or even become identical with it.

4.2 Model of the photon

In the time-dependent state of wave phenomena, equations (7)–(10) yield the following results:

- The Standard Model corresponds to a vanishing right-hand member of (1), and leads to the set of Hertz equations with a vanishing electric field divergence. In its turn, this gives rise to a vanishing photon spin as obtained from (23) and its quantized equivalent [5, 14]. Due to RQED theory there is on the other hand a photon model based on the extended relativistic forms of equations (6), (21) and (22), leading to a nonzero spin and an associated nonzero but very small rest mass [5, 14]. Thereby the spin of a photon wave packet does not merely have to be assumed in general terms, but becomes *deduced*. The spin occurs at the expense of a small reduction of the phase and group velocities in the direction of propagation.
- The needle-like photon model represented by equations (25a) and (25b) contributes to the *understanding* of the photoelectric effect and of two-slit experiments, with their wave-particle dualism.

- The RQED theory on screw-shaped wave modes is consistent with observed *hollow* geometry of corks-crew-shaped light beams [5].
- The nonzero electric field divergence and its intrinsic electric charges of alternating polarity also contributes to the understanding of electron-positron *pair formation* from an electrically neutral and energetic photon.

5 Conclusions

The present revised quantum electrodynamic theory includes the results of earlier fundamental theories, such as that by Dirac on electron spin and antimatter, and that by Higgs on massive elementary particles. It could thus be put into doubt whether the theory by Higgs becomes necessary for explaining the particle rest masses. In addition, the present theory leads to new results beyond those available from these and other so far established fundamental theories, as well as from the Standard Model in general.

Submitted on: July 29, 2013 / Accepted on: August 05, 2013

References

1. Schiff L. I. Quantum Mechanics. McGraw-Hill Book Comp.,Inc., New York-Toronto-London, 1949, Ch.IV, Sec.13; Ch.XIV.
2. Casimir H. B. G. On the attraction between two perfectly conducting plates. *Proc.K.Ned.Akad.Wet.*,1948, v. 51, 793–795.
3. Lamoreaux S. K. Demonstration of the Casimir force in the 0.6 to 6 μm range. *Phys.Rev.Letters*, 1997, v. 78 (1), 5–8.
4. Abbot L. The mystery of the cosmological constant. *Scientific American*, 1988, v. 258 (5), 106–113.
5. Lehnert B. Revised Quantum Electrodynamics. In Contemporary Fundamental Physics, Edited by V.V.Dvoeglazov. Nova Science Publishers, Inc., New York, 2013.
6. Morse P. M. and Feshbach H. Methods of Theoretical Physics. McGraw-Hill Book Comp., Inc., New York-Toronto-London, 1953, Part I, p.260.
7. Higgs P. W. Spontaneous symmetry breaking without massless bosons. *Physical Review*, 1966, v. 145, 1156–1168.
8. Lehnert B. A way to revised quantum electrodynamics. *Progress in Physics*, 2012, v. 2, April 2012, 21–27.
9. Heitler W. The Quantum Theory of Radiation. Third Edition, Oxford, Clarendon Press, 1954, Appendix, Sec.3, p.409.
10. Aad G. et al., ATLAS Collaboration, Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Phys.Lett*, 2012, v. B716, 1–29.
11. Chatrchyan S. et al., CMS Collaboration. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Phys.Lett*, 2012, v. B716, 30–61.
12. Lehnert B. Higgs-like particle due to revised quantum electrodynamics. P*rogress in Physics*, 2013, v. 3, 31–32.
13. Lehnert B. Intrinsic charges and the strong force. *Progress in Physics*, 2013, v. 3, 17–20.
14. Lehnert B. Qn angular momentum and rest mass of the photon. *Journal of Plasma Physics*, Memorial volume dedicated to Padma K. Shukla, 2013(in press).

A Complete Relativity Theory Predicts with Precision the Neutrino Velocities Reported by OPERA, MINOS, and ICARUS

Ramzi Suleiman
University of Haifa, Haifa 31509, Israel. E-mail: suleiman@psy.haifa.ac.il

The present paper utilizes the recently proposed Complete Relativity Theory (CR) for the prediction of neutrino velocity in a prototypical neutrino velocity experiment. The derived expression for the relative difference of the neutrino velocity with respect to the velocity of light is a function of the anticipation time δt, the traveled distance D and the light velocity c, measured on Earth. It is independent neither on the traveling particle type nor on its energy level. With regard to fast neutrinos it is shown that the derived equation predicts with precision the results reported by OPERA, MINOS, and ICARUS. Since CR postulates that all physical entities, including the velocity of light, are relativistic entities, it follows that even though the results of the aforementioned experiments fail to support the neutrino superluminality claim, their precise prediction based on a theory that diametrically opposes SR, provides strong evidence for the inadequacy of SR in accounting for the dynamics of quasi-luminal particles. The aforementioned notwithstanding, a direct calculation of SR's predictions for the above mentioned studies yields grossly incorrect results.

1 Introduction

The findings of several high energy experiments conducted by MINOS, OPERA, ICARUS and other collaborations suggest that neutrinos travel at super-luminal or quasi-luminal velocities, e.g. [1–6]. The possibility of quasi-luminal neutrinos has been also confirmed by cosmological observations, see, e.g. [7, 8]. Among all experimental findings, the one that attracted most interest was the result reported in 2011 by OPERA [1], which (ostensibly) indicated that neutrinos have travelled faster than light. The reported anticipation time was $\delta t = 60.7 \pm 6.9\,(\text{stat.}) \pm 7.4\,(\text{sys.})$ ns and the relative neutrino velocity was $\frac{v_n - c}{c} = (5.1 \pm 2.9) \times 10^{-5}$. Many physicists have described the possibility that OPERA may have broken the limit of light-velocity as one of the greatest discoveries in particle physics, provided that it is replicated by an independent group, and CERN's Research Director announced in a press conference that "If this measurement is confirmed, it might change our view of physics" [9].

Within few months, numerous papers were written, proposing that OPERA's experimental design and/or measurements were flowed, or suggesting various explanations that accord with standard theories, see, e.g. [10–20]. Soon after, the ICARUS collaboration reported a null result, which contradicted OPERA's superluminal one [3]. The anticipation time measured by ICARUS was $0.3 \pm 4.0\,(\text{stat}) \pm 9.0$ (sys.) ns, which is one order of magnitude lower than the result reported by OPERA [1]. The following events witnessed the discovery of hardware malfunctions which resulted in measurement error and the publication of a corrected null result [5].

Theoretically, the possibility of superluminal particles has been treated within the framework of General Relativity by A. Zelmanov's theory of "physically observable quantities" [21, 22]. Other models which entertain the possibility to construct theories in which neutrinos travel faster than photons have recently been proposed, e.g. [20, 23].

Although many questions pertaining to the neutrino superluminality issue remain open to theoretical inquiry, the general stance among physicists contends that for the time being both superluminality and subluminality of neutrinos cannot be dismissed by existing data, and that more investigation of this issue is needed [23, 24]. The common view, which I shall refute hereafter, contends that the null result based on data aggregation from existing experiments, is consistent with Special Relativity and with the limits put on Lorentz violations, e.g. [12, 15, 24, 25].

Here I shall show that for three experiments conducted by MINUS, OPERA, and ICARUS, Special Relativity (SR) yields *grossly incorrect results* and that an expression for $\frac{v_n - c}{c}$ derived on basis of Complete Relativity Theory (CR), detailed in [26] in this volume, yields *precise predictions* for the three aforementioned experiments.

The reminder of the paper is organized as follows: Section 2 details a derivation of $\frac{v_n - c}{c}$ based on SR, and demonstrates that it yields grossly incorrect predictions for all the discussed experiments. Section 3 provides a brief description of CR, and utilizes the one-way time transformation for deriving an expression for $\frac{v_n - c}{c}$ in a typical quasi-luminal neutrino experiment. The derived expression is then used to make precise predictions for the results reported by the above mentioned studies. Section 4 ends with concluding remarks.

2 Special Relativity predictions

In general, all neutrino-velocity experiments utilized the same technology. Thus, for the sake of convenience and without loss of generality, I analyze the one implemented by OPERA shown schematically in Fig. 1.

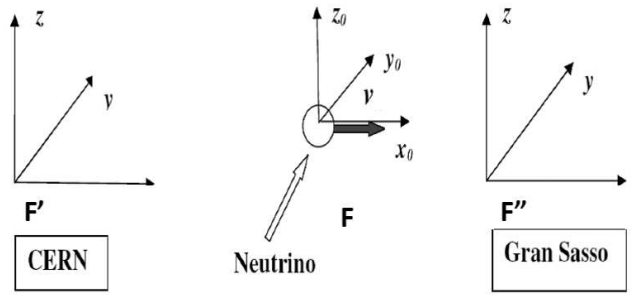

Fig. 1: The OPERA Setup.

From the perspective of Special Relativity (SR), the start and end laboratories F' and F'' are stationed in one frame of reference. The time dilation predicted by SR is given by:

$$\Delta_{SR} = T''_{G.Sasso} - T'_{CERN} = \frac{1}{\sqrt{1-\left(\frac{v_n}{c}\right)^2}}\, T. \qquad (1)$$

Where Δ_{SR} is the time difference between the start and end points, v_n is the neutrino's velocity, c is the velocity of light as it is measured on earth ($c = 299792.458$ km/sec) and T is the rest time at the neutrino's frame of reference F given by:

$$T = \frac{D}{v_n}. \qquad (2)$$

Where D is the distance between the source of the neutrino beam and the end point detector. Substituting the value of T in Eq. 1 we obtain:

$$\Delta_{SR} = \frac{1}{\sqrt{1-\left(\frac{v_n}{c}\right)^2}} \frac{D}{v_n}. \qquad (3)$$

For an early neutrino arrival time (δt) with respect to light photons we get:

$$\Delta_{SR} = \frac{D}{c} - \delta t. \qquad (4)$$

Substituting the value of Δ_{SR} from Eq. (3) in Eq. (4) and solving for $\frac{v_n}{c}$ we obtain:

$$\frac{v_n}{c} = \pm \sqrt{\frac{1}{2}\left(1 + \sqrt{1 - \frac{4}{1-\left(\frac{c\delta t}{D}\right)^2}}\right)}. \qquad (5)$$

For the result reported by ICARUS 2011: $\delta t = (0.3 \pm \pm 4.0\,(\text{stat}) \pm 9.0\,(\text{sys.})$ and $D = 674.385$ km. Substituting in Eq. 4 we get:

$$\frac{v_n}{c} \approx \pm(0.86603 + 0.5i) \qquad (6)$$

And,

$$\frac{c - v_n}{c} = \pm(-0.13397 + 0.5i). \qquad (7)$$

Calculations of SR's prediction of $\frac{c-v_n}{c}$ for the results reported by MINOS and OPERA (not reported here) yield similar (incorrect) results.

3 Complete Relativity predictions

Complete Relativity Theory (CR) rests on two postulates:

1. The magnitudes of all physical entities, as measured by an observer, depend on the relative motion of the observer with respect to the rest frame of the measured entities.
2. All translations of information from one frame of reference to another are carried by light or electromagnetic waves of equal velocity.

It should be stressed that the first postulate applies to all measured entities, *including the velocity of light*. CR treats the velocity of light as a relativistic quantity and not as an invariant one as postulated by SR. The derivations of CR's time, distance, mass-density and energy transformations are detailed elsewhere in this volume [26].

The derivation of a theoretical expression for $\frac{v_n - c}{c}$ in a typical superluminal neutrino experiment requires only the one-way time transformation. Viewed in the framework of CR, the experimental setup depicted in Fig. 1 includes *three* frames of reference: F' at CERN, F'' at Gran Sasso and F, the neutrino rest frame. F is *departing* from F' with velocity v_n and *approaching* F'' with velocity $-v_n$. F' and F'' are at rest relative to each other. According to CR [26], the time transformation for the one-way travel is given by:

$$\frac{t}{t_0} = \frac{1}{1 - \frac{v}{c}} \qquad (8)$$

Thus, we can write:

$$T'_{CERN} = \frac{1}{1 - \frac{v_n}{c}}\, T. \qquad (9)$$

Where v_n is the velocity of the neutrino relative to CERN's frame of reference F'.

Since the neutrino travelled *towards* Gran Sasso, applying the time transformation to F'' yields:

$$T''_{G.Sasso} = \frac{1}{1-\left(\frac{-v_n}{c}\right)}\, T = \frac{1}{1 + \frac{v_n}{c}}\, T. \qquad (10)$$

Experiment	Neutrino Anticipation Time (δt)	Experimental $\frac{v_n - c}{c}$	Theoretical $\frac{v_n - c}{c}$
MINOS $D = 734298.6$ m	$\left(126 \pm 32\,(\text{stat.}) \pm 6\,(\text{sys.})\right)$ ns	$(5.1 \pm 2.9)\,(\text{stat}) \times 10^{-5}$	$\mathbf{5.14 \times 10^{-5}}$
OPERA 2012 (corrected result) $D = 730085$ m	$\left(6.5 \pm 7.4\,(\text{stat.}) \begin{smallmatrix}+9.2\\-6.8\end{smallmatrix} (\text{sys.})\right)$ ns	$\left(2.7 \pm 3.1\,(\text{stat.}) \begin{smallmatrix}+3.8\\-2.8\end{smallmatrix} (\text{sys.})\right) \times 10^{-6}$	$\mathbf{2.67 \times 10^{-6}}$
ICARUS 2012 $D = 730478.56$ m	$\left(0.10 \pm 0.67\,(\text{stat.}) \pm 2.39\,(\text{sys.})\right)$ ns	$\left(0.4 \pm 2.8\,(\text{stat.}) \pm 9.8\,(\text{sys.})\right) \times 10^{-7}$	$\mathbf{0.41 \times 10^{-7}}$

Table 1: Experimental results and theoretical predictions for three superluminal neutrino experiments.

The time difference between CERN and Gran Sasso's could be written as:

$$D_t = T''_{G.Sasso} - T'_{CERN} = \\ = \left[\frac{1}{1+\frac{v_n}{c}} - \frac{1}{1-\frac{v_n}{c}}\right] T = -\frac{2\frac{v_n}{c}}{1-\left(\frac{v_n}{c}\right)^2} T. \quad (11)$$

Substituting the value of T in Eq. 11 we obtain:

$$D_t = -\frac{2\frac{v_n}{c}}{1-\left(\frac{v_n}{c}\right)^2} \frac{D}{v_n}. \quad (12)$$

For an early neutrino arrival time of δt with respect to the velocity of light we can write:

$$D_t = -\frac{2\frac{v_n}{c}}{1-\left(\frac{v_n}{c}\right)^2} \frac{D}{v_n} = \frac{D}{c} - \delta t. \quad (13)$$

Where $\frac{D}{c}$ is the light time arrival from CERN to Grand Sasso. Solving Eq. 13 for $\frac{v_n}{c}$ yields:

$$\frac{v_n}{c} = \sqrt{\frac{2}{1-\frac{c\delta t}{D}} - 1}\,. \quad (14)$$

Or,

$$\frac{v_n - c}{c} = \sqrt{\frac{2}{1-\frac{c\delta t}{D}} - 1} - 1. \quad (15)$$

Predictions

For the OPERA *corrected* result [2]

$$\delta t = \left(6.5 \pm 7.4\,(\text{stat.}) \begin{smallmatrix}+9.2\\-6.8\end{smallmatrix} (\text{sys.})\right) \text{ns}$$

and $D = 730.085$ km. Substituting in Eq. 15 we get:

$$\frac{v_n - c}{c} = \sqrt{\frac{2}{1-\frac{299792.458 \times 6.5 \times 10^{-9}}{730.085}} - 1} - 1 \approx \\ \approx \mathbf{2.67 \times 10^{-6}}. \quad (16)$$

Which is identical to the reported result of:

$$\frac{v_n - c}{c}\,(\text{Exp.}) = \left(2.7 \pm 3.1\,(\text{stat.}) \begin{smallmatrix}+3.8\\-2.8\end{smallmatrix} (\text{sys.})\right) \times 10^{-6}. \quad (17)$$

Equation 15 was also used to calculate theoretical predictions for the results reported by ICARUS [4] and MINOS [5]. The results are summarized in Table 1, which depicts all three experimental results against the corresponding theoretical predictions.

As could be seen in the table, *CR* yields accurate predictions for *all* three experimental results, *including the null ones*.

4 Concluding remarks

In this article I applied a recently proposed Complete Relativity Theory (CR) to analyze the neutrino travel in a typical neutrino-velocity experiment. CR treats all physical entities, including light velocity, as relativistic entities. Accordingly the measured velocity of light depends on the direction of the light propagation vector, relative to the laboratory. In terms of relative time, the start point laboratory (e.g., at CERN) will measure *time dilation*, whereas the end point laboratory (e.g., at Gran Sasso) will measure *time contraction*. It is important to note that the CR-based model presented in section 3 is independent of the particle type and its energy level. For the prediction of $\frac{v_n - c}{c}$ only the anticipation time δt and distance D between the start and end points are required [see Eq. 15].

The analysis brought above shows that CR predicts with near precision all the relative neutrino velocities $\frac{v_n - c}{c}$ obtained in recent neutrino-velocity experiments. In contrast, SR's predictions for all the discussed findings yields grossly incorrect results. What becomes clear from the analysis brought above

is that a breakdown of Einstein's SR does not require that the neutrino velocity exceeds the velocity of light.

Upon the announcement of the first null result, the leader of ICARUS collaboration leader was quoted by the press saying that had they found 60 nanoseconds, he would have sent a bottle of champagne to OPERA, and that instead, he suspects that he "will be toasting Einstein" [31]. The analysis presented in the present paper suggests that the news about rescuing SR were premature, and that it makes more sense to keep the champagne in the frigidaire.

Submitted on May 14, 2013 / Accepted on August 19, 2013

References

1. Adam T., et al. Measurement of the neutrino velocity with the OPERA detector in the CNGS beam. (OPERA Collaboration), 2011. arXiv:1109.4897.
2. Adam T., et al. Measurement of the neutrino velocity with the OPERA detector in the CNGS beam (OPERA Collaboration). *Journal of High Energy Physics*, 2012, v. 10, 093. arXiv:1109.4897.
3. Antonello M. et al. Measurement of the neutrino velocity with the ICARUS detector at the CNGS beam. *Physics Letters B*, 2012, v. 713 (1), 17–22.
4. Antonello M. et al. Precision measurement of the neutrino velocity with the ICARUS detector in the CNGS beam. *Journal of High Energy Physics*, 2012, 49. arXiv:1208.2629.
5. Adamson P. et al. Measurement of neutrino velocity with the MINOS detectors and NuMI neutrino beam. (MINOS Collaboration). *Physical Review D*, 2007, v. 76 (7), 2005–2012. arXiv:0706.0437.
6. Kalbfleisch G. R., Baggett N., Fowler C., Alspector J. Experimental comparison of neutrino, antineutrino, and muon velocities. *Physical Review Letters*, 1979, v. 43, 1361–1364.
7. Stodolsky L. The speed of light and the speed of neutrinos. *Physics Letters B*, 1988, v. 201 (3), 353–354.
8. Longo M. J. Tests of relativity from SN1987A. *Physical Review D*, 1987, v. 36, 3276–3277.
9. See announcement by said CERN Research Director, Prof. Sergio Bertolucci at: http://press.web.cern.ch/press/PressReleases/Releases 20 11/PR19.11E.html.
10. Garcia-Islas J. M. A very simple solution to the OPERA neutrino velocity problem. 2011. arXiv:1110.5866.
11. Mbelek J-P. Special relativity is consistent with the opera measurements of the neutrino "velocity". The Smithsonian/NASA Astrophysics Data System, 2011. arXiv:1110.4095.
12. Torrealba R. OPERA superluminal neutrinos explained by spontaneous emission and stimulated absorption, 2011. arXiv:1112.0815.
13. Cowsik R., Nussinov S., Sarkar U. Superluminal neutrinos at OPERA confront pion decay kinematics. *Physical Review Letters*, 2011, v. 107 (25), 251801.
14. Zichichi A. Time-shift in the OPERA set-up: Proof against superluminal neutrinos without the need of knowing the CERN-LNGS distance. *International Journal of Modern Physics*, 2012, v. 27 (18), 1230017. arXiv:1206.2840.
15. Schechter J., Shahid N. Neutrinos with velocities greater than c? *Physical Review D*, 2012, v. 85, 093008, 2012. arXiv:1201.4374.
16. Giudicea G. F., Sibiryakovb S., Strumiac A. Interpreting OPERA results on superluminal neutrino. *Nuclear Physics B*, 2012, v. 861 (1), 2012, 1–16.
17. Oda I. Subluminal OPERA neutrinos. *International Journal of Modern Physics A*, 2012, v. 27 (7), 1250033. arXiv:1112.5793.
18. Bo-Qiang, M. The phantom of the OPERA: superluminal neutrinos. *Modern Physics Letters A*, 2012, v. 27 (4). arXiv:1111.7050.
19. Hebecker A., Knochel A. The price of neutrino superluminality continues to rise. *Physics Letters B*, 2012, v. 715 (1-3), 116–120. arXiv:1111.6579.
20. Mu-Lin Y., Neng-Chao X., Huang W., Hu S. Superluminal neutrinos from special relativity with de sitter space-time symmetry. *Mod. Phys. Lett. A*, 2012, v. 27 (14), 1250076. arXiv:1111.4532.
21. Zelmanov A. On the relativistic theory of an anisotropic inhomogeneous universe (translated from the Russian manuscript, published in 1957). *The Abraham Zelmanov Journal*, 2008, v. 2008 (1), 33–63.
22. Rabounski D., Borissova L. Particles here and beyond the mirror. American Research Press, 2012.
23. Alexandre J., Ellis J., Mavromatos N. E. On the possibility of superluminal neutrino propagation. *Physics Letters B*, 2012, v. 706 (4-5), 456–461.
24. Linglia Z., Maa B-Q. Neutrino speed anomaly as signal of Lorentz violation. *Astroparticle Physics*, 2013, v. 44, 24–27.
25. Klinkhamer F. R. Superluminal neutrino, flavor, and relativity. *Physical Review D*, 2012, v. 85, 016011.
26. Suleiman R. The dark side revealed: a complete relativity theory predicts the content of the universe. *Progress in Physics*, 2013, v. 4, 34–40.
27. O'Connor J. J., Roberston E. F. Christian Andreas Doppler-MacTutor history of mathematics archive. University of St. Andrews, 1998.
28. Michelson A., Morley E. W. Relative motion of earth and luminiferous ether. *American Journal of Science*, 1887, v. 34, 333–345.
29. Piccard A., Stahel E. L'expérience de Michelson, réalisée en ballon libre. *Comptes Rendus*, 1926, v. 183 (7), 420–421.
30. Illingworth K. K. A Repetition of the Michelson-Morley experiment using Kennedy's refinement. *Physical Review*, 1927, v. 30 (5), 692–696.
31. Brumfiel G. Neutrinos not faster than light ICARUS experiment contradicts controversial claim. *Nature News*, 2012. http://www.nature.com/news/neutrinos-not-faster-than-light-1.10249.

Nanotechnology Quantum Detectors for Gravitational Waves: Adelaide to London Correlations Observed

Reginald T. Cahill

School of Chemical and Physical Sciences, Flinders University, Adelaide 5001, Australia. E-mail: reg.cahill@flinders.edu.au

The discovery of the nanotechnology zener diode quantum detector effect for gravitational waves is reported, based upon the quantum to classical transition being induced by dynamical 3-space fluctuations. Gravitational waves were detected by way of waveform correlations between time measurement in two Digital Storage Oscilloscopes, revealing time delays of 13 to 20 seconds over 24 hrs for Adelaide to London travel, varying as the earth rotates. The speed and direction were found, for January 1, 2013, to be 512 km/s, RA = 4.8 hrs, Dec = 83 deg S. This velocity agrees with previous detections using different techniques, such as the NASA spacecraft Earth-flyby Doppler shifts, which found 491 km/s, RA = 5.2 hrs, Dec = 80 deg S, for December 8, 1992. Consequently it was realised that nanotechnology zener diode quantum detectors have been operating, for different reasons, for some 15 years, and are known as RNGs (Random Number Generators) or REGs (Random Event Generators). The discovery herein reveals that they are not random. Correlations between data from a REG in Perth and a REG in London gave the speed and direction, for January 1, 2013, to be 528 km/s, RA = 5.3 hrs, Dec = 81 deg S. We also report highly correlated current fluctuations from collocated zener diode circuits. The GCP REG network constitutes an international gravitational wave detector network, with currently some 60 REGs operating, and with records going back to 1998. These detectors permit the study of dynamical 3-space structure, and also apparent anomalous scattering of the waves when passing deeper into the earth, solar flares, coronal mass ejections, earthquakes, and correlations with fluctuations in various rate processes such as nuclear decays. The quantum to classical transition is shown to be caused by 3-space dynamics, and so challenges the standard interpretation of probabilities in quantum theory.

1 Introduction

The speed and direction of gravitational waves have been directly measured via waveform time delays from detectors located in Adelaide and London, and separately from Perth and London. The Adelaide to London correlations were detected utilising the discovery that so-called "clock jitter" in two digital storage oscilloscopes (DSO) is actually correlated, with the London signal delayed relative to the Adelaide signal by 13 to 20 seconds, depending on sidereal time, so that at least part of the clock jitter is actually induced by passing gravitational waves. Subsequently similar correlations were discovered in Random Event Generator (REG) correlated data. These detect the quantum to classical transition for electrons tunnelling through a barrier in a tunnel diode, a nanotechnology device. According to the standard interpretation of quantum theory such electron current fluctuations should be completely random, which is why such devices are also known as hardware Random Number Generators (RNG), and have a variety of applications assuming such randomness.

These discoveries make the detection and study of gravitational waves particularly simple, and easily extend to a network of detectors, and for the REG technique an international network of such detectors has existed since 1998, and so that data is an extremely valuable to the characterisation of the gravitational wave effect, and also other phenomena which appear to be induced by more extreme fluctuations. Correlations of the gravitational wave forms permit determination of the speed and direction of space, which agrees with results from NASA Earth-flyby Doppler shift data, and with the 1925/26 Dayton Miller Mt.Wilson gas-mode Michelson interferometer data. The correlation data also reveals two new phenomena: a speed-up when the waves pass deeper into the earth, and a wave reverberation effect. For collocated zener diodes the current fluctuations are highly correlated, with no time delay effects, as expected. The quantum to classical transition is thus shown to be caused by 3-space dynamics, and so challenges the standard interpretation of probabilities in quantum theory.

2 Classical physics gravitational wave detectors

Classical gravitational wave detectors have employed a number of physical effects and designs: gas-mode Michelson interferometers, optical fibre Michelson interferometers, RF coaxial cable travel time differential measurements, and more compact RF coaxial cable – optical fibre measurements, spacecraft Earth-flyby Doppler effects, and dual RF coaxial cable travel time measurements [1,2]. All of these techniques utilise light or EMR anisotropy speed effect in a single device.

The key issue with such devices is that they are single-site devices, and require a calibration theory, which depends upon an assumed theory. For example the sensitivity of a Michelson interferometer, as indicated by the travel time difference between the two arms, and detected by means of fringe shifts as the detector is rotated, is given by

$$\Delta t = k^2 \frac{L v_p^2}{c^3} \cos(2(\theta - \psi)) \qquad (1)$$

where L is the arm length, v_P is the speed projected onto the plane of the interferometer, and the angles measure the rotation effect, see [1]. Eqn.(1) is applied to the data in conjunction with terms accounting for the inclined mirrors and temperature drift effects [1]. The critical factor k^2 is the calibration constant. With a gas present in the light path, with refractive index n, $k^2 \approx n^2 - 1$ to a good approximation. Results from two gas-mode Michelson interferometer experiments are shown in Fig. 1. The results reveal significant turbulence, which has been identified as gravitational waves, and much greater in magnitude than expected. Michelson and Morley in the 1st such experiment in 1887 assumed that $k^2 = 1$, whereas with air present, $n = 1.00029$, giving $k^2 \approx 0.0006$, and so much less sensitive than assumed. Note that a vacuum-mode Michelson interferometer has $k^2 = 0$, and so completely insensitive to gravitational waves.

A recent gravitational wave experiment used differential travel time measurements in a dual RF coaxial cable array [2]. This technique relies upon the absence of Fresnel drag in RF coaxial cables, at least for low RF frequencies (~10 MHz). The results agree with those form the Miller gas-mode Michelson interferometer, and from the NASA flyby Doppler shift data. The fluctuations were again observed to be a ~20% effect.

The interpretation of the magnitude of the detected effects in these classical detector experiments all rely upon some calibration theory, and there has always been confusion. Fortunately spacecraft flyby Doppler shift analysis does not suffer from such problems, and has indeed confirmed the results from the classical detectors. We now report the discovery that nanotechnology quantum detectors respond to the fluctuations of the passing space, and when the data from two well-separated detectors is subject to a correlation analysis of the two local waveforms the average speed and direction of the passing space is revealed, together with significant wave/turbulence effects. This technique gives an absolute measurement of travel times.

3 Quantum gravitational wave detectors

When extending the Dual RF Coaxial Cable Detector experiment to include one located in London, in addition to that located in Adelaide, an analysis of the measured DSO internal noise in each identically setup instrument was undertaken, when the extensive RF coaxial cable array was replaced by

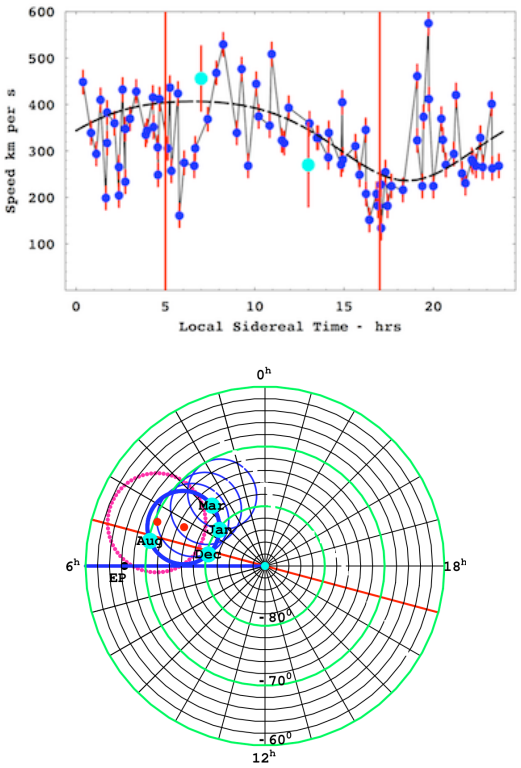

Fig. 1: Top: Speeds v_P, of the space velocity **v** projected onto the horizontal plane of the Miller gas-mode Michelson interferometer located atop Mt.Wilson, plotted against local sidereal, for a composite day, with data collected over a number of days in September 1925, see [1]. The data shows considerable fluctuations, from hour to hour, and also day to day, as this is a composite day. The dashed curve shows the non-fluctuating best-fit variation over one day, as the earth rotates, causing the projection onto the plane of the interferometer of the velocity of the average direction of the space flow to change. The maximum projected speed from the curve is 417 km/s, corresponding to a speed of 453 km/s, with a RA of ~5 hrs, which is very close to results reported herein. The Cassini flyby Doppler shift data in August 1999 gives a RA = 5.2 hrs [1]. The green data points, with error bars, at 7 hrs and 13 hrs, are from the Michelson-Morley 1887 data. The ~20% speed fluctuations are seen to be much larger than statistically determined errors, revealing the presence of turbulence in the space flow, i.e gravitational waves. Bottom: South celestial pole region. The dot (red) at RA = 4.3^h, Dec = 75°S, and with speed 486 km/s, is the direction of motion of the solar system through space determined from NASA spacecraft earth-flyby Doppler shifts [1], revealing the EM radiation speed anisotropy. The thick (blue) circle centred on this direction is the observed velocity direction for different days of the year, caused by earth orbital motion and sun 3-space inflow. The corresponding results from the Miller gas-mode interferometer are shown by 2nd dot (red) and its aberration circle (red dots). For December 8, 1992, the velocity is RA = 5.2^h, Dec = 80°S, speed 491 km/s, see Table 2 of [1]. The thinner blue aberration circles relate to determination of earth 3-space inflow speed, see [1].

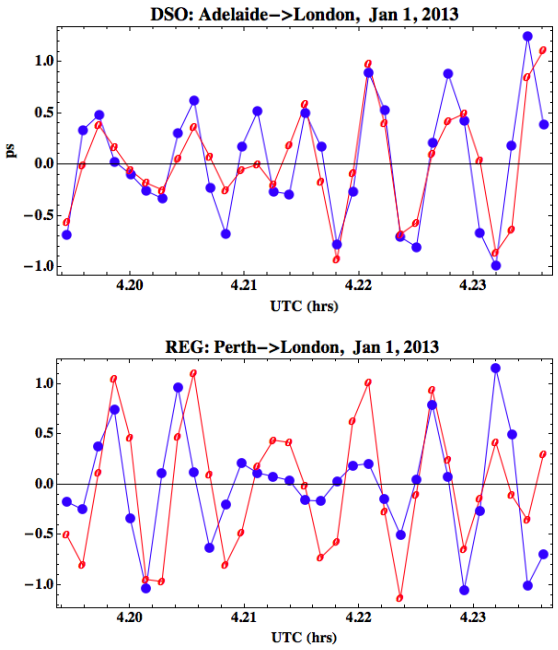

Fig. 2: Correlations in band-passed Adelaide-London DSO data (top) and Perth (Australia)-London REG data (bottom), for January 1, 2013, with London data (red, open dots) advanced by 15 s in both cases, over the same 200 s time interval. The data points are at 5 s intervals. In-phase correlations from collocated Zener Diode Detectors are shown in Fig. 7. The REG data was recorded every 1 s, and has been averaged to 5 s intervals for ease of comparison with DSO data. The data shows a quasi-periodicity of ~20 s, related to the reverberation effect [3]. The UTC time at all detectors was determined using internet timing applications, which have ms precision.

short leads. This was intended to determine the S/N ratio for the joint Adelaide-London experiment. Surprisingly the internal noise was found to be correlated, with the noise in the London DSO being some 13 to 20 seconds behind the Adelaide DSO* noise, see Fig. 2. The correlation data had a phase that tracked sidereal time, meaning that the average direction was approximately fixed wrt the galaxy, but with extensive fluctuations as well from the gravitational wave/turbulence effect, that had been seen in all previous experiments. The explanation for this DSO effect was not possible as the DSO is a complex instruments, and which component was responding to the passing space fluctuations could not be determined. But the correlation analysis did demonstrate that not all of the internal noise in the DSO was being caused solely by some random process intrinsic to the instrument. Subsequent experiments, below, now suggest that there are zener diodes within the time difference measurements hardware within the DSO.

The travel time delay $\tau(t)$ was determined by computing the correlation function

$$C(\tau,t) = \int_{t-T}^{t+T} dt' S_1(t'-\tau/2) S_2[t'+\tau/2] e^{-a(t'-t)^2} \qquad (2)$$

for the two detector signals $S_1(t)$ and $S_2(t)$. Here $2T = 200s$ is the time interval used, about UTC time t. The gaussian term ensures the absence of end-effects. Maximising $C(\tau,t)$ wrt τ gives $\tau(t)$ - the delay time vs UTC t, and plotted in Figs. 3 and 4, where the data has been binned into 1hr time intervals, and the rms also shown. The speed and direction, over a 24hr period, was determined by fitting the time delay data using

$$\tau = \frac{\mathbf{R}\cdot\mathbf{v}}{\mathbf{v}^2}, \qquad (3)$$

where $\mathbf{R}$ is the Adelaide-London spatial separation vector, and $\mathbf{v}(\theta,\delta)$ is the 3-space velocity vector, parametrised by a speed, RA and Declination. This expression assumes a plane wave form for the gravitational waves. The $\tau(t)$ delay times show large fluctuations, corresponding to fluctuations in speed and/or direction, as also seen in data in Fig. 1, and also a quasi-periodicity, as seen in Fig. 2. Then only minimal travel times, 10 s $< \tau <$ 22 s, were retained. Correlations, as shown in Fig. 2, are not always evident, and then the correlation function $C(\tau,t)$ has a low value. Only $\tau(t)$ data from high values of the correlation function were used. The absence of correlations at all times is expected as the London detector is not directly "downstream" of the Adelaide detector, and so a fractal structure to space, possessing a spatial inhomogeneity, bars continuous correlations, and as well the wave structure will evolve during the travel time. Fig. 2 shows examples of significant correlations in phase and amplitude between all four detectors, but with some mismatches. The approximate travel time of 15 s in Fig. 2 at ~4.2 hrs UTC is also apparent in Fig. 3, with the top figure showing the discovery of the correlations from the two DSO separated by a distance $R \approx 12160$ km. That the internal "noise" in these DSO is correlated is a major discovery.

There are much simpler devices that were discovered to also display time delayed correlations over large distances: these are the Random Number Generators (RNG) or Random Event Generators (REG). There are various designs available from manufacturers, and all claim that these devices manifest hardware random quantum processes, as they involve the quantum to classical transition when a measurements, say, of the quantum tunnelling of electrons through a nanotechnology potential barrier, ~10 nm thickness, is measured by a classical/macroscopic system. According to the standard interpretation of the quantum theory, the collapse of the electron wave function to one side or the other of the barrier, after the tunnelling produces a component on each side, is purely a random event, internal to the quantum system. However this interpretation had never been tested experimentally. Guided by the results from the DSO correlated-noise effect, the data

*LeCroy WaveRunner 6051A DSOs were used.

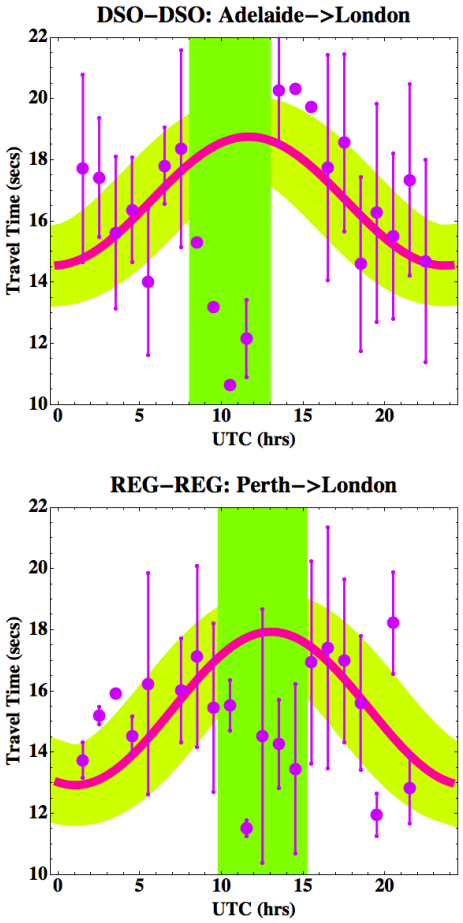

Fig. 3: Travel times from DSO-DSO Adelaide-London data (top), and REG-REG Perth-London data (bottom) from correlation analysis using (2). The data in each 1 hr interval has been binned, and the average and rms shown. The thick (red line) shows best fit to data using plane wave travel time predictor, (3), but after excluding those data points between 8 and 13 hrs UTC (top) and 10 and 15 hrs UTC (bottom), indicated by vertical band. Those data points are not consistent with the plane wave modelling, and suggest a scattering process when the waves pass deeper into the earth, see Fig. 5. The Perth-London phase is retarded wrt Adelaide-London phase by ~1.5 hrs, consistent with Perth being 1.5 hrs west of Adelaide. The Adelaide-London data gives speed = 512 km/s, RA = 4.8 hrs, Dec = 83°S, and the Perth-London data gives speed = 528 km/s, RA = 5.3 hrs, Dec = 81°S. The broad band tracking the best fit line is for ±1 sec fluctuations, corresponding to speed fluctuation of ±17 km/s. Actual fluctuations are larger than this, as 1st observed by Michelson-Morley and by Miller, see Fig. 1.

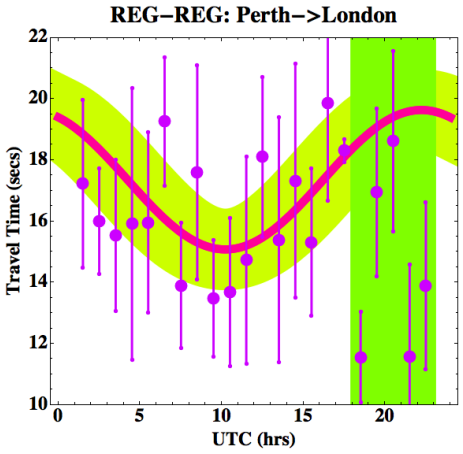

Fig. 4: Travel times from REG-REG Perth-London data for August 1, 2012. The data in each 1 hr interval has been binned, and the average and rms shown. The thick (red line) shows best fit to data using plane wave travel time predictor, (3), but after excluding those data points between 18 and 23 hrs UTC, indicated by vertical band. Those data points are not consistent with the plane wave modelling. This data gives speed = 471 km/s, RA = 4.4 hrs, Dec = 82°S. The change in phase of the maximum of the data, from UTC = 22±2 hr, for August 1, 2012, to UTC = 12±2 hr for January 2013 (Fig. 3), but with essentially the same RA, illustrates the sidereal effect: the average direction of the space flow is fixed wrt to the stars, apart from the earth-orbit aberration effect, Fig. 1.

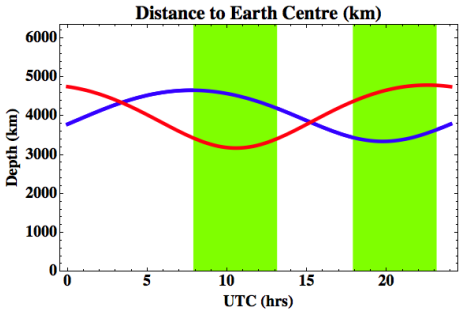

Fig. 5: Given measured space velocity, plots show maximum earth penetration depth of space detected by London detectors for Adelaide→London, Jan1, 2013 (red) and Perth→London, August 1, 2012 (blue), revealing that the anomalous scattering occurs when deeper depths are "traversed". The vertical shadings correspond to those in Fig. 3 (top) and Fig. 4.

from two REGs, located in Perth and London, was examined. The data* showed the same correlation effect as observed in the DSO experiments, see Figs. 2–4. However REGs typically employ a XOR gate that produces integer valued outputs with a predetermined statistical form. To study the zener diode tunnelling currents without XOR gate intervention two collocated zener diode circuits were used to detect highly correlated tunnelling currents, Figs. 6 and 7. When the detectors are separated by ~0.5 m, phase differences ~μs were observed and dependent on relative orientation. So this zener diode circuit forms a very simple and cheap nanotechnology quantum detector for gravitational waves.

*The data is from the GCP international network: `http://teilhard.global-mind.org/`

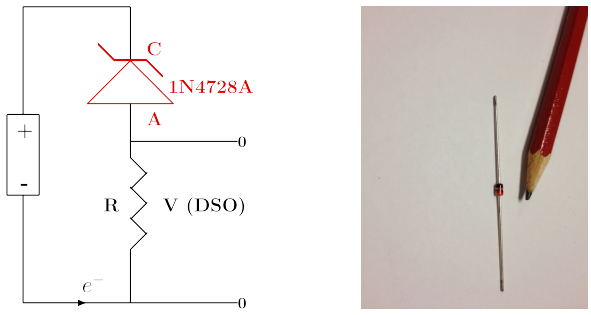

Fig. 6: Left: Circuit of Zener Diode Gravitational Wave Detector, showing 1.5 V AA battery, 1N4728A zener diode operating in reverse bias mode, and having a Zener voltage of 3.3 V, and resistor R = 10 kΩ. Voltage V across resistor is measured and used to determine the space driven fluctuating tunnelling current through the zener diode. Correlated currents from two collocated detectors are shown in Fig. 7. This design avoids data degradation from the XOR gate in commercial REGs. Right: Photo of zener diode showing size in comparison to pencil tip. The zener diode costs $0.5.

4 Dynamical 3-space gravitational waves

It is necessary to give some background to the interpretation of reported correlations as gravitational waves. Experiments and theory have suggested that space is a dynamical system:

$$\nabla\cdot\left(\frac{\partial \mathbf{v}}{\partial t}+(\mathbf{v}\cdot\nabla)\mathbf{v}\right)+\frac{5\alpha}{4}\left((trD)^2-tr(D^2)\right)+ \delta^2\nabla^2\left((trD)^2-tr(D^2)\right)+...=-4\pi G\rho, \quad (4)$$

where $D_{ij}=\partial v_i/\partial x_j$ and $\rho(\mathbf{r},\mathbf{t})$ is the usual matter density. This entails a velocity field $\mathbf{v}(\mathbf{r},t)$ describing the motion of a structured 3-space relative to an observers frame of reference. This easily follows from writing Newtonian gravity in terms of a velocity field, which then permits additional terms, with coefficients α and δ. This field and its fluctuations has been repeatedly detected over some 125 years. The 1st term, the Newtonian gravity term, involves the Euler 3-space constituent acceleration, while the $\alpha-$ and $\delta-$ terms contain higher order derivative terms and describe the self interaction of space. Laboratory, geophysical and astronomical data suggest that α is the fine structure constant $\approx 1/137$, while δ appears to be a very small but non-zero Planck-like length. The emergence of gravity arises from the unique coupling of quantum theory to the 3-space, which determines the 'gravitational' acceleration of quantum matter as a quantum wave refraction effect,

$$\mathbf{g}=\frac{\partial \mathbf{v}}{\partial \mathbf{t}}+(\mathbf{v}\cdot\nabla)\mathbf{v}+(\nabla\times\mathbf{v})\mathbf{v_R}-\frac{\mathbf{v_R}}{1-\frac{\mathbf{v_R^2}}{\mathbf{c^2}}}\frac{\mathbf{1}}{\mathbf{2}}\frac{\mathbf{d}}{\mathbf{dt}}\left(\frac{\mathbf{v_R^2}}{\mathbf{c^2}}\right) \quad (5)$$

where $\mathbf{v_R}=\mathbf{v_0}-\mathbf{v}$ is the velocity of quantum matter relative to the local space. The 1st two terms are the Euler 3-space acceleration, the 2nd term explains the Lense-Thirring effect when the vorticity is non-zero, and the last term explains the precession of planetary orbits. Neglecting relativistic effects (4) and (5) give

$$\nabla\cdot\mathbf{g}=-\mathbf{4\pi G\rho}-\mathbf{4\pi G\rho_{DM}}, \quad (6)$$

where ρ_{DM} is the α and δ term, describing a 3-space self-interaction effects, with the α term explaining the so-called 'dark matter' effects. The spatial dynamics is non-local and exhibits instantaneous effects, which points to the universe being highly connected, consistent with the deeper pre-space *Process Physics* [6]. Historically this was first noticed by Newton who called it action-at-a-distance. This shows a high degree of non-locality and non-linearity, and in particular that the behaviour of both ρ_{DM} and ρ manifest at a distance irrespective of the dynamics of the intervening space. A key implication of (6) is that observed fluctuations in $\mathbf{v}(\mathbf{r},t)$ can only generate gravitational effects via the ρ_{DM} processes. So the velocity field is more fundamental than the Newtonian gravitational acceleration field. Although not presented herein significant fluctuations in $\mathbf{v}(\mathbf{r},t)$ were observed to be correlated with solar flares, coronal mass ejections, and earthquakes. These effects suggest that the 11 year solar cycle is caused by galactic-scale larger than normal 3-space fluctuations. The delay of several days between major fluctuations and solar flares implies that the new 3-space/gravitational wave detectors may be used as an early warning system for such solar flares.

One consequence of the non-linearity of (4) is that fluctuations in $\mathbf{v}(\mathbf{r},t)$ develop reverberations [3], which are clearly apparent in the data in Fig. 2. Another implication suggested by the data is that when the 3-space fluctuations penetrate the earth the non-linearity cause the 3-space waveforms to manifest at a distance, without propagating through the intervening space, resulting in an apparent speed-up, as manifestly evident in the data – an effect that had to be taken into account in the analysis based upon a normal plane-wave like propagation, as indicated by the vertical bands in Figs. 3 and 4. The data from numerous experiments clearly shows that the so-called "gravitational waves" have observed properties very different from those commonly assumed.

5 Probability in Quantum Theory

The conventional quantum theories all have the generic form $i\hbar\partial\psi/\partial t=H\psi$, differing only by the configuration space on which ψ is based, and the Hamiltonian. The interpretation has been, as proposed by Born, that $|\psi|^2$ is the probability density for the location of a particle, which is assumed to exist apart from ψ. However missing from this generic unitary time evolution for ψ is (i) the existence of a dynamical 3-space, as distinct from the usual frame of reference, and which leads to gravity as am emergent phenomenon, and (ii) the existence of terms which model the localisation of ψ in space by a classical detector of quantum waves [5]. In [6, p. 40], it was

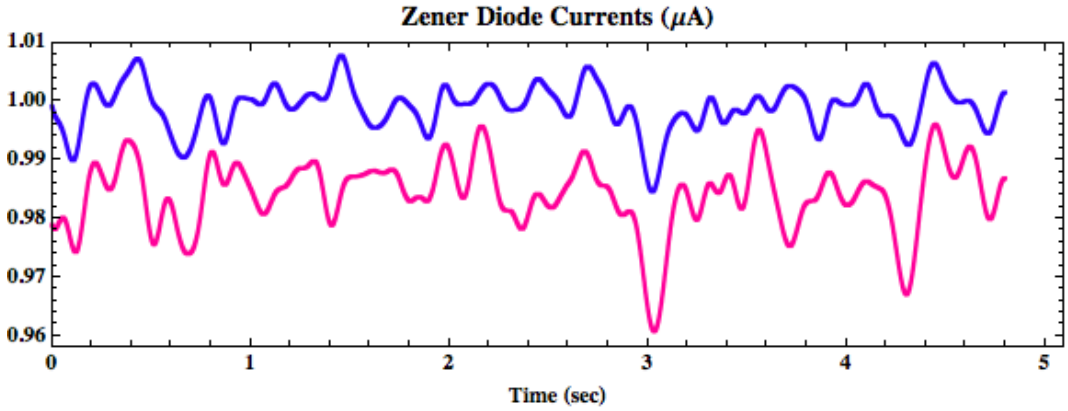

Fig. 7: Zener Diode tunnelling currents over 5 sec interval, showing higher time resolution than in Fig. 2. Band pass filter was used to remove higher frequencies. Plots have been displaced vertically for ease of viewing. The two zener diode circuits were collocated with the zener diodes separated by ~30 mm. Highly correlated currents are observed, demonstrating that the tunnelling currents are not random, as required by the conventional interpretation of quantum theory, and as 1st discovered in the Adelaide-London correlations.

argued that emergent classicality, including the ψ localisation effects, are caused by fluctuations in the 3-space. This and the present results would amount to the discovery that reality is fundamentally only quantum waves embedded in a quantum foam space, and that the classical world is an emergent macroscopic phenomenon: our reality is induced by the nature of 3-space fluctuations.

6 Conclusions

We have reported the discovery of the quantum detection of gravitational waves, showing correlations between well separated locations, that permitted the absolute determination of the 3-space velocity of some 500 km/s, in agreement with the speed and direction from a number of previous analyses, including in particular the NASA spacecraft earth-flyby Doppler shift effect. This discovery enables a very simple and cheap nanotechnology zener diode quantum gravitational wave detection technology, which will permit the study of various associated phenomena, such as solar flares, coronal mass ejections, earthquakes, eclipse effects, moon location effects, non-Poisson fluctuations in radioactivity [4], and other rate processes, and variations in radioactive decay rates related to distance of the earth from the Sun, as the 3-space fluctuations are enhanced by proximity to the sun.

Acknowledgements

This report is from the Flinders University Gravitational Wave Detector Project, funded by an Australian Research Council Discovery Grant: *Development and Study of a New Theory of Gravity*. Thanks to Professor Igor Bray for making this work possible. Thanks to GCP and its director Dr. Roger Nelson for the availability of extensive and valuable data from the REG international network:
`http://teilhard.global-mind.org/`
Thanks also to Dr. David Brotherton-Ratcliffe for providing the London DSO data and for several useful discussions.

Submitted on: September 01, 2013 / Accepted on: September 04, 2013

References

1. Cahill R.T. Combining NASA/JPL one-way optical-fiber light-speed data with spacecraft Earth-flyby Doppler-shift data to characterise 3-space flow. *Progress in Physics*, 2009, issue 4, 50–64.
2. Cahill R.T. Characterisation of low frequency gravitational waves from dual RF coaxial-cable detector: fractal textured dynamical 3-space. *Progress in Physics*, 2012, issue 3, 3–10.
3. Cahill R.T. and Deane S.T. Dynamical 3-space gravitational waves: reverberation effects. *Progress in Physics*, 2013, issue 2, 9–11.
4. Shnoll S.E. Cosmophysical Factors in Stochastic Processes. American Research Press, Rehoboth, New Mexico, USA, 2012.
5. Percival I. Quantum State Diffusion. Cambridge University Press, 1998.
6. Cahill R.T. Process Physics: From Information Theory to Quantum Space and Matter. Nova Science Pub., New York, 2005.

Atomic Weights Confirm Bipolar Model of Oscillations in a Chain System

Andreas Ries
Universidade Federal de Pernambuco, Centro de Tecnologia e Geociências, Laboratório de Dispositivos e Nanoestruturas,
Rua Acadêmico Hélio Ramos s/n, 50740-330 Recife – PE, Brazil
E-mail: andreasries@yahoo.com

We apply the bipolar model of oscillations in a chain system to the data set of standard atomic weights. 90% of these masses could be reproduced by this model and were expressed in continued fraction form, where all numerators are Euler's number and the sum of the free link and all partial denominators yields zero. All outliers were either radioactive or polynuclidic elements whose isotopic compositions as found in samples on Earth might not be fully representative for the mean values when considering samples from all parts of the universe.

1 Introduction

In several previous papers we applied the model of oscillations in a chain system to various systems such as the solar system [1], excited electronic states of atoms [2], the electron density in the Hydrogen atom [3], and more recently to the mass distribution of elementary particles [4].

Initially, this model was founded by Müller in three fundamental articles [5–7] and assumes that all protons in the universe are oscillators, coupled through the physical vacuum. As a consequence we can consider (in the most simplest case) a chain of equal harmonic proton oscillators with an associated logarithmic spectrum of eigenfrequencies which can be expressed through continued fractions. In that way, every mass is interpreted as a proton resonance state and expressed in continued fraction form.

Recently, a bipolar version of this model was proposed for the description of elementary particles [4], because the traditional version could not reproduce their masses in a fully satisfactory way. The idea of bipolarity postulates that the fundamental spectrum of proton resonances has an opposite, an anti-oscillation or inverted oscillation spectrum, and this is the spectrum of electron resonances.

Mathematically, two opposite oscillation states are characterized through equal continued fraction representations, but with the difference that in one case all denominators, the free link and the phase shift have been multiplied by (-1). From the analysis of elementary particle masses it was suggested to express the electron mass as a proton resonance and the proton mass as an electron resonance through the following equations (e is Euler's number):

$$\ln \frac{m_{electron}}{m_{proton}} = p_p + (-6) + \cfrac{e}{12 + \cfrac{e}{-6}}$$

and

$$\ln \frac{m_{proton}}{m_{electron}} = p_e + 6 + \cfrac{e}{(-12) + \cfrac{e}{6}}.$$

Numerically, p_p was found to be ≈ -1.75 [4]; for these phase shifts must hold p_p = -p_e.

In this article we show that the relative atomic masses can be reproduced by almost the same the bipolar model. The only parameter that must be adjusted is the phase shift (from $|p| \approx 1.75$ to $|p| \approx 1.79$) and this is a very minor change.

2 Data sources and computational details

The standard atomic weights, including the proton and electron reference masses were taken from the web-site of the National Institute of Standards (NIST) and were expressed in the atomic mass unit u. The following abbreviations and conventions for the numerical analysis hold:

The atomic masses are transformed into a continued fraction according to the equations

$$\ln \frac{m}{m_{electron}} = p_e + S, \qquad \ln \frac{m}{m_{proton}} = p_p + S,$$

where p is the phase shift (it must hold p_p = -p_e) and S is the continued fraction (e is Euler's number)

$$S = n_0 + \cfrac{e}{n_1 + \cfrac{e}{n_2 + \cfrac{e}{n_3 + ...}}}. \qquad (1)$$

The numerical value of the phase shift p is initially unknown and must be adjusted in such a way that the largest possible amount of atomic weights can be expressed through a continued fraction.

The continued fraction representation $p+S$ is abbreviated as $[p; n_0 \mid n_1, n_2, n_3, \dots]$, where the free link n_0 is allowed to be $0, \pm3, \pm6, \pm9 \dots$ and all partial denominators n_i can take the values $e+1, -e-1, \pm6, \pm9, \pm12 \dots$. In the tables these abbreviations were marked with P or E, in order to indicate proton or electron resonance states.

The absolute value of the difference between the atomic weight given by NIST and the atomic weight calculated from the associated continued fraction representation is defined as numerical error and listed in the tables.

An atomic weight is considered as an outlier when the corresponding continued fraction representation provides a mass value outside the interval "atomic mass ± standard deviation".

Table 1: Continued fraction representations of the 20 most accurately determined atomic weights (Helium and the set of 19 mononuclidic non-radioactive elements), x = 1.7918229 is the phase shift, SD = standard deviation.

Element symbol	Mass ± SD [u]	Continued fraction representation(s)	Numerical error [u]
He	$4.002602 \pm 2.0 \times 10^{-6}$	P [-x; 3 | 15, e+1, -6, e+1, 33, (6, -e-1, -e-1, -51)]	1.2×10^{-9}
Be	$9.0121822 \pm 4.0 \times 10^{-7}$	P [-x; 3 | e+1, -e-1, e+1, -e-1, -e-1, e+1, -6, -6, 12, (-9, 6)]	4.3×10^{-8}
F	$18.99840322 \pm 7.0 \times 10^{-8}$	E [x; 9 | -9, e+1, -e-1, e+1, -e-1, -1680, (1680)]	2.6×10^{-9}
Na	$22.9897692809 \pm 2.9 \times 10^{-9}$	E [x; 9 | -18, -9, -9, -e-1, e+1, -33, 12, (48)]	7.9×10^{-10}
Al	$26.98153863 \pm 1.2 \times 10^{-7}$	P [0; 3 | 9, 6, 18, -6, -6, e+1, 9, (-e-1, 33)]	1.2×10^{-9}
P	$30.97376163 \pm 2.0 \times 10^{-7}$	P [-x; 6 | -e-1, 12, -e-1, 6, -15, e+1, 6, (-15, e+1)] E [x; 9 | 18, 15, -18, 18, (-42)]	1.6×10^{-7} 1.1×10^{-7}
Sc	$44.9559119 \pm 9.0 \times 10^{-7}$	P [0; 3 | e+1, -9, 9, -e-1, e+1, -21, e+1, (-e-1, 18, -e-1)] E [0; 12 | -e-1, -12, e+1, -6, 126, e+1, -e-1, -120]	4.5×10^{-7} 9.1×10^{-8}
Mn	$54.9380451 \pm 7.0 \times 10^{-7}$	P [0; 3 | e+1, -e-1, e+1, -e-1, e+1, -e-1, 6, 18, e+1, -e-1, e+1, -e-1, (-27)] P [-x; 6 | -12, -e-1, e+1, -e-1, e+1, -6, 111, (-99)] E [0; 12 | -6, 6, e+1, -24, e+1, -e-1, e+1, -6, (18, -e-1, -e-1)] E [x; 9 | e+1, 63, -e-1, 6, -9, -18, (-51)]	2.2×10^{-7} 4.5×10^{-8} 1.2×10^{-8} 3.6×10^{-10}
Co	$58.9331950 \pm 7.0 \times 10^{-7}$	E [0; 12 | -6, -6, e+1, -e-1, e+1, -e-1, e+1, -e-1, 162, (-162)] E [x; 9 | e+1, -9, -6, -e-1, e+1, -24, -9, (-e-1, 39)]	5.7×10^{-8} 4.4×10^{-9}
As	$74.9215965 \pm 2.0 \times 10^{-6}$	P [-x; 6 | 27, -30, -27, (24)]	2.3×10^{-7}
Y	$88.9058483 \pm 2.7 \times 10^{-6}$	P [-x; 6 | 9, e+1, -e-1, e+1, -e-1, -102, -e-1, (87, e+1)]	5.1×10^{-7}
Nb	$92.9063781 \pm 2.6 \times 10^{-6}$	P [-x; 6 | 9, -6, -e-1, 21, e+1, -e-1, (27, e+1, -57)] E [0; 12 | 69, -e-1, -6, 18, -e-1, (-93, e+1, e+1)]	2.2×10^{-8} 6.9×10^{-7}
Rh	$102.905504 \pm 3.0 \times 10^{-6}$	P [-x; 6 | 6, 6, -6, e+1, -e-1, -6, -6, -e-1, (e+1)]	9.1×10^{-7}
I	$126.904473 \pm 4.0 \times 10^{-6}$	P [-x; 6 | e+1, e+1, e+1, -93, 6, -e-1, e+1, (-e-1, 81, -e-1, -e-1)]	3.6×10^{-7}
Cs	$132.905451933 \pm 2.4 \times 10^{-8}$	E [0; 12 | 6, e+1, -6, 99, e+1, 6, -6, (-111, -e-1, -e-1)]	1.7×10^{-8}
Pr	$140.9076528 \pm 2.6 \times 10^{-6}$	P [-x; 6 | e+1, -330, -e-1, 6, -e-1, e+1, (-12, 330)] E [0; 12 | 6, -63, -e-1, e+1, 9, -12, (48)]	4.4×10^{-8} 2.5×10^{-7}
Tb	$158.9253468 \pm 2.7 \times 10^{-6}$	P [-x; 6 | e+1, -6, e+1, -e-1, -6, 525, (-519, -e-1)]	5.8×10^{-8}
Ho	$164.9303221 \pm 2.7 \times 10^{-6}$	P [0; 6 | -e-1, e+1, 18, e+1, -6, 6, 75, (-99, -e-1)]	5.2×10^{-9}
Tm	$168.9342133 \pm 2.7 \times 10^{-6}$	P [0; 6 | -e-1, e+1, e+1, 6, 6, 6, 12, (15, -e-1, -51)]	3.2×10^{-9}
Au	$196.9665687 \pm 6.0 \times 10^{-7}$	P [0; 6 | -e-1, -78, e+1, e+1, 6, -e-1, e+1, (15, -e-1, 51)] E [0; 12 | e+1, -9, -e-1, -e-1, -15, e+1, 6, -e-1, e+1, -e-1, (e+1, 6)]	7.1×10^{-9} 4.6×10^{-7}

3 Results and discussion

It can be easily verified that the standard Müller model with the phase shifts p = 0 and p = 1.5 does not apply at all to the relative atomic weights, while the bipolar model with phase shifts of approximately ± 1.75 (as used in a previous study [4]) produces around 30% outliers. When working with the complete data set, varying the phase shift does not lead to a clear result. In that case we obtain a wealth of slightly different phase shifts, all providing a quite similar number of outliers and a similar sum of squared residuals (sum of squared numerical errors).

In order to arrive at a conclusion, the data set of 84 atomic masses was divided into two parts. The first part is composed of the element Helium (two stable isotopes, but still very low standard deviation) and the set of 19 non-radioactive mononuclidic elements. Here, the maximum measurement error is 2.7×10^{-6} u. The second part consists of the remaining elements; their standard deviations vary from $\sim 10^{-5}$ to 0.1 u (Pb) due to isotopic variations found in samples taken at different locations on Earth.

It is fact that the "mean atomic mass" of a mononuclidic element is everywhere in the universe exactly the same, while we would expect some variations in the atomic masses of polynuclidic elements when analyzing rock samples obtained from different galaxies. It is reasonable to assume that the conditions during the formation of the chemical elements were subjected to variations throughout the universe.

Therefore we give priority to the atomic masses of the mononuclidic chemical elements and only the first part of the data set has been analyzed thoroughly. The phase shift was adjusted in such a way that (a) the number of outliers, and (b) the sum of squared residuals are minimized. This leads to a phase shift of ≈ 1.79 (exact value is 1.7918229) which is close to ln(6). Table 1 lists these atomic masses together with

Table 2: Continued fraction representations of the atomic weights of polynuclidic elements, from H to Kr (except He), x = 1.7918229 is the phase shift, SD = standard deviation.

Element symbol	Mass ± SD [u]	Continued fraction representation(s)	Numerical error [u]
H	$1.00794 \pm 7 \times 10^{-5}$	P [0; 0 | 4128, (-4128)] E [x; 6 | -9, e+1, 6, -e-1, e+1, -e-1, e+1, (-e-1, 9, -12)]	2.5×10^{-8} 5.9×10^{-5}
Li	$6.941 \pm 2 \times 10^{-3}$	P [-x; 3 | e+1, 57, (e+1, -e-1, -e-1, -60)] E [0; 9 | 6, 27, (-42)]	5.5×10^{-4} 4.9×10^{-4}
B	$10.811 \pm 7 \times 10^{-3}$	P [0; 3 | -e-1, -e-1, -e-1, -9, (e+1, 6, e+1, e+1)] E [0; 9 | e+1, -e-1, -6, -e-1, (21, -24, e+1)] E [x; 9 | -e-1, e+1, 24, (-33)]	2.0×10^{-4} 5.5×10^{-5} 3.5×10^{-3}
C	$12.0107 \pm 8 \times 10^{-4}$	E [0; 9 | e+1, -e-1, e+1, -e-1, e+1, e+1, -6, (e+1, -9, 6, -e-1, -e-1, -e-1)] E [x; 9 | -e-1, 9, -9, -e-1, e+1, (-9, e+1)]	1.4×10^{-6} 6.8×10^{-4}
N	$14.0067 \pm 2 \times 10^{-4}$	E [x; 9 | -e-1, -6, e+1, e+1, -e-1, 9, (-12)]	6.2×10^{-5}
O	$15.9994 \pm 3 \times 10^{-4}$	P [0; 3 | -12, 6, 6, -e-1, (-e-1, 6, -9, e+1, e+1)]	9.0×10^{-7}
Ne	$20.1797 \pm 6 \times 10^{-4}$	P [0; 3 | -1056, (1053)] E [x; 9 | -9, -e-1, 42, (-42, e+1)]	2.3×10^{-5} 1.6×10^{-5}
Mg	$24.3050 \pm 6 \times 10^{-4}$	P [0; 3 | 15, -15, (-138, 135)] E [x; 9 | -30, e+1, -18, (-e-1, 39)]	8.4×10^{-9} 8.7×10^{-5}
Si	$28.0855 \pm 3 \times 10^{-4}$	P [0; 3 | 9, -e-1, -27, -e-1, e+1, (15, e+1)] P [-x; 6 | -e-1, e+1, e+1, e+1, -6, 18, (-e-1, -18, -e-1)]	2.7×10^{-4} 2.1×10^{-5}
S	$32.065 \pm 5 \times 10^{-3}$	P [0; 3 | 6, -27, -e-1, (e+1, 18)] P [-x; 6 | -e-1, 33, (-33, -6, e+1)] E [x; 9 | 15, -12, (6, -18)]	3.1×10^{-4} 1.4×10^{-5} 2.7×10^{-4}
Cl	$35.453 \pm 2 \times 10^{-3}$	E [0; 12 | -e-1, e+1, -12, -e-1, 6, (-6, e+1)]	9.1×10^{-5}
Ar	$39.948 \pm 1 \times 10^{-3}$	P [0; 3 | e+1, 9, e+1, -9, e+1, (-e-1, 6, -e-1, -e-1, -9)] P [-x; 6 | -6, e+1, -6, e+1, -12, (18, -e-1, -e-1)] E [0; 12 | -e-1, 9, -e-1, e+1, -e-1, (-21, e+1, e+1)] E [x; 9 | 6, e+1, 348, (-e-1, -363)]	4.3×10^{-7} 6.8×10^{-4} 3.9×10^{-4} 2.0×10^{-5}
K [Outlier]	$39.0983 \pm 1 \times 10^{-4}$	P [0; 3 | e+1, 6, e+1, 9, -e-1, e+1, -e-1, e+1, -e-1]	6.3×10^{-4}
Ca	$40.078 \pm 4 \times 10^{-3}$	P [-x; 6 | -6, e+1, -6, -e-1, 6] E [0; 12 | -e-1, 9, -e-1, -e-1, -e-1, (-21, e+1, e+1, e+1, e+1)] E [x; 9 | 6, e+1, 9, (-24, -e-1)]	1.1×10^{-3} 1.3×10^{-4} 8.0×10^{-4}
Ti	$40.078 \pm 1 \times 10^{-3}$	E [0; 12 | -e-1, -e-1, -6, e+1, 45, (-51, e+1)] E [x; 9 | e+1, e+1, -e-1, 9, (-23202, -e-1, 23184)]	1.2×10^{-5} 6.8×10^{-12}
V [Outlier]	$50.9415 \pm 1 \times 10^{-4}$	P [0; 3 | e+1, -e-1, 12, e+1, -e-1, e+1, -e-1, e+1, -e-1]	4.7×10^{-4}
Cr	$51.9961 \pm 6 \times 10^{-4}$	P [0; 3 | e+1, -e-1, 6, -9, -12, (12)] E [0; 12 | -6, e+1, -e-1, e+1, -15, e+1, (-e-1, 9, -e-1)]	1.8×10^{-5} 3.4×10^{-5}
Fe	$55.845 \pm 2 \times 10^{-3}$	P [-x; 6 | -15, e+1, -e-1, -30, (39)] E [0; 9 | e+1, -69, 6, (54, -e-1)]	1.6×10^{-4} 2.4×10^{-4}
Ni	$58.6934 \pm 4 \times 10^{-4}$	P [-x; 6 | -18, -e-1, e+1, -e-1, e+1, -e-1, (12, e+1)] E [0; 12 | -6, -6, 9, -e-1, e+1, (-9)]	2.9×10^{-4} 3.9×10^{-4}
Cu	$63.546 \pm 3 \times 10^{-3}$	P [-x; 6 | -42, -e-1, -15, (51, e+1)] E [x; 9 | e+1, -e-1, -e-1, e+1, -75, (66)]	8.5×10^{-5} 7.5×10^{-5}
Zn	$65.38 \pm 2 \times 10^{-2}$	P [-x; 6 | -78, e+1, (-e-1, 72)] E [0; 12 | -9, 9, e+1, (-e-1, -12)] E [x; 9 | e+1, -e-1, -12, e+1, (-21, 24, -e-1)]	3.4×10^{-3} 4.4×10^{-3} 9.8×10^{-5}
Ga	$69.723 \pm 1 \times 10^{-3}$	P [-x; 6 | 93, 6, e+1, (-105, -e-1)] E [x; 9 | e+1, -e-1, e+1, 6, e+1, -e-1, e+1, -e-1, (-e-1, -15)]	3.3×10^{-5} 4.5×10^{-6}
Ge	$72.64 \pm 1 \times 10^{-2}$	P [-x; 6 | 39, -12, (-33)]	1.1×10^{-3}
Se	$78.96 \pm 3 \times 10^{-2}$	P [-x; 6 | 18, -9, (-15)]	2.1×10^{-3}
Br	$79.904 \pm 1 \times 10^{-3}$	E [0; 12 | -24, -6, 6, (12)]	4.4×10^{-4}
Kr	$83.798 \pm 2 \times 10^{-3}$	P [-x; 6 | 12, e+1, -18, (6, -e-1, -6)] E [0; 12 | -42, -e-1, e+1, e+1, (-e-1, 30)]	3.4×10^{-5} 2.3×10^{-4}

Table 3: Continued fraction representations of the atomic weights of polynuclidic elements, from Rb to Os, x = 1.7918229 is the phase shift, SD = standard deviation.

Element symbol	Mass ± SD [u]	Continued fraction representation(s)	Numerical error [u]
Rb	$85.4678 \pm 3 \times 10^{-4}$	P [-x; 6 | 12, -9, 6, -6, -e-1, (-9, e+1)] E [0; 12 | -63, e+1, -15, (66, -e-1)]	2.3×10^{-4} 2.1×10^{-4}
Sr	$87.62 \pm 1 \times 10^{-2}$	E [0; 12 | -144, (132)]	5.8×10^{-3}
Zr	$91.224 \pm 2 \times 10^{-3}$	P [-x; 6 | 9, 21, e+1, (-36, -e-1)] E [0; 12 | 126, 6, (-144)]	2.8×10^{-4} 4.6×10^{-5}
Mo [Outlier]	$95.96 \pm 2 \times 10^{-2}$	E [0; 12 | 39, -e-1, e+1, -e-1, e+1]	5.7×10^{-2}
Ru	$101.07 \pm 2 \times 10^{-2}$	P [-x; 6 | 6, e+1, -9, -e-1, (6, -9)] E [0; 12 | 21, e+1, -e-1, (12, -45)]	3.5×10^{-3} 8.6×10^{-4}
Pd	$106.42 \pm 1 \times 10^{-2}$	P [-x; 6 | 6, 189, (6, -207)] E [0; 12 | 15, 6, -6, (-27)]	1.5×10^{-5} 8.2×10^{-3}
Ag [Outlier]	$107.8682 \pm 2 \times 10^{-4}$	E [0; 12 | 15, -e-1, -e-1, -6, e+1, -e-1, e+1, -e-1]	3.3×10^{-4}
Cd [Outlier]	$112.411 \pm 8 \times 10^{-3}$	E [0; 12 | 12, -15, e+1, -e-1, e+1, -e-1]	1.1×10^{-2}
In	$114.818 \pm 3 \times 10^{-3}$	P [-x; 6 | 6, -e-1, 6, -e-1, 6, e+1, (-24, e+1)]	2.5×10^{-3}
Sn	$118.710 \pm 7 \times 10^{-3}$	E [0; 12 | 9, 6, -e-1, e+1, -e-1, e+1, (-27)]	1.5×10^{-5}
Sb	$121.760 \pm 1 \times 10^{-3}$	P [-x; 6 | e+1, e+1, -e-1, 30, 6, (-42, -e-1)] E [0; 12 | 9, -12, e+1, e+1, -e-1, e+1, (-e-1, -9, -e-1)]	8.6×10^{-5} 6.3×10^{-4}
Te [Outlier]	$127.60 \pm 3 \times 10^{-2}$	P [-x; 6 | e+1, e+1, e+1, -e-1, e+1, -e-1, e+1]	7.2×10^{-2}
Xe [Outlier]	$131.293 \pm 6 \times 10^{-3}$	P [-x; 6 | e+1, 6, e+1, -e-1, e+1, -e-1, e+1, -e-1]	1.1×10^{-2}
Ba	$137.327 \pm 7 \times 10^{-3}$	P [-x; 6 | e+1, 21, 6, (-33, -e-1)] E [0; 12 | 6, 9, -9, e+1, (-18, -e-1)]	4.9×10^{-3} 8.9×10^{-4}
La	$138.90547 \pm 7 \times 10^{-5}$	P [-x; 6 | e+1, 42, -e-1, -e-1, -6, (-42, e+1)]	3.1×10^{-5}
Ce	$140.116 \pm 1 \times 10^{-3}$	P [-x; 6 | e+1, 132, 6, (-e-1, -144)] E [0; 12 | 6, 84, e+1, -e-1 (-102)]	5.8×10^{-4} 8.1×10^{-4}
Nd	$144.242 \pm 3 \times 10^{-3}$	E [0; 12 | 6, -9, e+1, -6, (9, -e-1, -12)]	8.9×10^{-4}
Sm	$150.36 \pm 2 \times 10^{-2}$	P [0; 6 | -e-1, e+1, -e-1, e+1, -e-1, -e-1, (-18, 12, e+1, e+1)] P [-x; 6 | e+1, -9, 12, (-e-1, -9)]	1.9×10^{-4} 2.8×10^{-3}
Eu	$151.964 \pm 1 \times 10^{-3}$	P [0; 6 | -e-1, e+1, -e-1, e+1, e+1, -6, (-69, -e-1, 69)] E [0; 12 | 6, -e-1, e+1, 6, e+1, -24, (-e-1)]	8.3×10^{-6} 8.7×10^{-5}
Gd	$157.25 \pm 3 \times 10^{-2}$	P [0; 6 | -e-1, e+1, -6, e+1, -e-1, e+1, (9, -e-1, -9)] P [-x; 6 | e+1, -6, 6, -12, (-e-1, 6)]	9.7×10^{-4} 1.2×10^{-3}
Dy	$162.500 \pm 1 \times 10^{-3}$	P [-x; 6 | e+1, -e-1, -e-1, -45, (e+1, 39)] E [0; 12 | e+1, e+1, -6, -e-1, -9, -e-1, e+1, (-e-1, 9, -6)]	5.0×10^{-4} 7.4×10^{-4}
Er	$167.259 \pm 3 \times 10^{-3}$	P [0; 6 | -e-1, e+1, 6, 6, e+1, (-18, -e-1)] P [-x; 6 | e+1, -e-1, -27, 30, (-9)] E [0; 12 | e+1, e+1, e+1, 63, (-e-1, -75, -e-1, -e-1)]	3.8×10^{-4} 1.3×10^{-4} 2.0×10^{-4}
Yb	$173.054 \pm 5 \times 10^{-3}$	P [0; 6 | -e-1, 6, -e-1, e+1, (222, e+1, -234)] P [-x; 6 | e+1, -e-1, 6, e+1, 9, e+1, (-e-1, -e-1, -21)] E [0; 12 | e+1, 6, e+1, -e-1, e+1, -e-1, -e-1, (-18)]	1.5×10^{-5} 8.6×10^{-4} 1.0×10^{-3}
Lu	$174.9668 \pm 1 \times 10^{-4}$	P [0; 6 | -e-1, 6, -6, 9, -e-1, e+1, e+1, (-15)] P [-x; 6 | e+1, -e-1, 6, -e-1, e+1, -6, -e-1, e+1, -e-1, (-6, e+1)]	1.4×10^{-5} 7.3×10^{-5}
Hf	$178.49 \pm 2 \times 10^{-2}$	P [-x; 6 | e+1, -e-1, e+1, -48, (-e-1, 42)]	1.9×10^{-3}
Ta	$180.94788 \pm 2 \times 10^{-5}$	E [0; 12 | e+1, 21, -39, -6, (-e-1, 12)]	1.1×10^{-5}
W	$183.84 \pm 1 \times 10^{-2}$	P [0; 6 | -e-1, 9, 9, -e-1, (e+1, -24, e+1)] P [-x; 6 | e+1, -e-1, e+1, -e-1, e+1, -e-1, 15, (-21)] E [0; 12 | e+1, 60, (-72, -e-1)]	3.4×10^{-3} 2.3×10^{-3} 1.8×10^{-3}
Re	$186.207 \pm 1 \times 10^{-3}$	P [0; 6 | -e-1, 12, -6, -12, (e+1)] E [0; 12 | e+1, -135, (123, -e-1)]	1.2×10^{-4} 5.8×10^{-4}
Os	$190.23 \pm 3 \times 10^{-2}$	E [0; 12 | e+1, -21, -e-1, (9)]	3.6×10^{-3}

Table 4: Continued fraction representations of the atomic weights of polynuclidic and radioactive elements, from Ir to U, x = 1.7918229 is the phase shift, SD = standard deviation.

Element symbol	Mass ± SD [u]	Continued fraction representation(s)	Numerical error [u]
Ir	$192.217 \pm 3 \times 10^{-3}$	E [0; 12 | e+1, -15, -e-1, -9, -e-1, (e+1, 12)]	1.1×10^{-3}
Pt	$195.084 \pm 9 \times 10^{-3}$	P [0; 6 | -e-1, 195, (-201, e+1)] E [0; 12 | e+1, -12, e+1, 15, (-e-1, -e-1, -15)]	1.9×10^{-3} 4.2×10^{-4}
Hg	$200.59 \pm 2 \times 10^{-2}$	P [0; 6 | -e-1, -21, 6, (e+1, 9)]	1.6×10^{-3}
Tl	$204.3833 \pm 2 \times 10^{-4}$	P [0; 6 | -e-1, -12, 6, -e-1, e+1, e+1] E [0; 12 | e+1, -6, -12, 6, e+1, -e-1, (-e-1)] E [x; 12 | -e-1, e+1, -e-1, -12, -27, (27, e+1)]	1.9×10^{-4} 5.4×10^{-5} 7.9×10^{-5}
Pb	207.2 ± 0.1	P [0; 6 | -e-1, -9, 6, (12, -15, e+1)] E [0; 12 | e+1, -6, 6, (27, -39, -e-1)] E [x; 12 | -e-1, e+1, -6, e+1, -e-1, e+1, -e-1 (-6)]	1.4×10^{-3} 8.4×10^{-5} 6.8×10^{-2}
Bi [Outlier]	$208.98040 \pm 1 \times 10^{-5}$	E [0; 12 | e+1, -6, e+1, -9, -e-1, -e-1, 6, e+1, -e-1, e+1, -e-1]	1.8×10^{-5}
Pa	$231.03588 \pm 2 \times 10^{-5}$	E [x; 12 | -e-1, 6, -6, -e-1, -e-1, -18, -6, (e+1, e+1, e+1, 12)]	1.2×10^{-6}
Th	$232.03806 \pm 2 \times 10^{-5}$	E [0; 12 | e+1, -e-1, e+1, e+1, -9, e+1, e+1, -9, e+1, -e-1 (6, -e-1, -e-1, -e-1, -e-1)] E [x; 12 | -e-1, 6, -12, -e-1, 6, 6, -e-1 (9, e+1, e+1, e+1, -27)]	1.7×10^{-5} 5.0×10^{-7}
U	$238.02891 \pm 3 \times 10^{-5}$	E [0; 12 | e+1, -e-1, e+1, -6, e+1, e+1, -e-1, -15, (-e-1, -e-1, 9)]	1.8×10^{-5}

the corresponding continued fraction representations and the numerical errors. As it can be seen, no outlier is present.

Moreover, many continued fractions show the effect of successively canceling denominators. For instance, the continued fraction representation for Be, as calculated by the computer is: P [-x; 3 | e+1, -e-1, e+1, -e-1, -e-1, e+1, -6, -6, 12, (-6, -360, ...)]. The denominators in brackets are not required to obtain a mass value inside the interval "atomic mass ± SD". Through a minimal manipulation, we obtain a zero sum of all denominators and the free link, without significantly changing the value of the fraction: P [-x; 3 | e+1, -e-1, e+1, -e-1, -e-1, e+1, -6, -6, 12, (-9, 6)]. As this procedure can be applied in a similar way to all elements, we demonstrate this and opted to express all continued fractions as a zero sum. Only redundant denominators (given in brackets) were manipulated to achieve the zero sums.

In a second step, the so-adjusted model was tested against the remaining 64 chemical elements. Only eight outliers were found (K, V, Mo, Ag, Cd, Te, Xe, Bi [radioactive]). Tables 2 to 4 show the results; for outliers, the best possible continued fraction is displayed (not as a zero sum), and it can be seen that in most cases the atomic mass is reproduced with a numerical error very little higher than the standard deviation.

4 Conclusions

The relative atomic masses are now the second data set that can be described by the bipolar model of oscillations in a chain system. In total, 10% outliers were found which might be attributed to the fact that the isotopic compositions of these outlier elements as found here on Earth are not good representatives for the true mean compositions when considering samples from distant parts of the universe.

Anyway, it is important to note that all mononuclidic elements can be described perfectly by this model.

Acknowledgments

The author greatly acknowledges the financial support from the Brazilian governmental funding agencies FACEPE and CNPq.

Submitted on September 11, 2013 / Accepted on September 15, 2013

References

1. Ries A. and Fook M.V.L. Application of the model of oscillations in a chain system to the solar system. *Progress in Physics*, 2011, v. 7(1), 103–111.
2. Ries A. and Fook M.V.L. Excited electronic states of atoms described by the model of oscillations in a chain system. *Progress in Physics*, 2011, v. 7(4), 20–24.
3. Ries A. The radial electron density in the Hydrogen atom and the model of oscillations in a chain system. *Progress in Physics*, 2012, v. 8(3), 29–34.
4. Ries A. A bipolar model of oscillations in a chain system for elementary particle masses. *Progress in Physics*, 2012, v. 8(4), 20–28.
5. Müller H. Fractal scaling Models of resonant oscillations in chain systems of harmonic oscillators. *Progress in Physics*, 2009, v. 5(2), 72–76.
6. Müller H. Fractal scaling models of natural oscillations in chain systems and the mass distribution of the celestial bodies in the solar system. *Progress in Physics*, 2010, v. 6(1), 62–66.
7. Müller H. Fractal scaling models of natural oscillations in chain systems and the mass distribution of particles. *Progress in Physics*, 2010, v. 6(3), 61–66.

Laboratory Instrument for Detecting the Earth's Time-Retarded Transverse Vector Potential

Joseph C. Hafele

Retired Physicist, Home Office: 618 S. 24th St., Laramie, WY, USA. E-mail: cahafele@bresnan.net

This article provides the basic design for a laboratory instrument that may detect the Earth's time-retarded transverse vector potential [Hafele J.C. *Zelm. Jour.*, 2012, v.5, 134]. The instrument is based on the compound pendulum used by N.A. Kozyrev to measure the change in weight of a suspended aircraft navigation gyroscope [Kozyrev N.A. *Zelm. Jour.*, 2012, v.5, 188]. If such an instrument is developed to measure the strength of the Earth's vector potential with a precision of about 1 part in 1000, the neoclassical causal theory can be worked backwards to calculate the speed of the Earth's gravitational field.

Introduction

A new causal version for Newtonian gravitational theory has been shown to explain exactly the six Earth flyby anomalies reported by NASA in 2008, and also explain exactly an overlooked lunar orbit anomaly [1, 2]. The new causal theory, which retains the traditional acausal radial component, requires in addition a small time-retarded transverse component for the Earth's gravitational field. The new transverse component is orthogonal to the traditional radial component and is directed along the east-west direction. It is well-known that the traditional radial component can be derived from the gradient of a scalar potential. However, the time-retarded transverse component can be derived only from the curl of a vector potential. The formula for the vector potential will be found by using Stoke's theorem. The resulting vector potential is directed along the north-south direction. The north-south component of the gravitational field is given by the time-derivative of the vector potential. By using an analogous Lorentz force law, it will be shown that a small time-dependent radial component is created by induction from the north-south gravitational field. This small induced radial component can slightly change the weight of a suspended gyroscope. By measuring the change in weight, the neoclassical causal theory can be worked backwards to deduce the strength of the vector potential, and thereby indirectly measure the speed of the Earth's gravitational field.

More than 60 years ago [3], N.A. Kozyrev used the causality principle to predict the need for a second universal velocity, one that is to be associated with rotational motion [4]. He designates c_2 as the speed for this second universal velocity. He developed a theory that suggests that the numerical value for c_2 should be related to the fine structure constant [5]. In electrostatic cgs units, the unit of electric charge is the statcoulomb.

The formula for the fine-structure constant, designated by α, in cgs electrostatic units, becomes [5]

$$\alpha = \frac{2\pi}{c}\frac{e^2}{h} \cong \frac{1}{137}, \qquad (1)$$

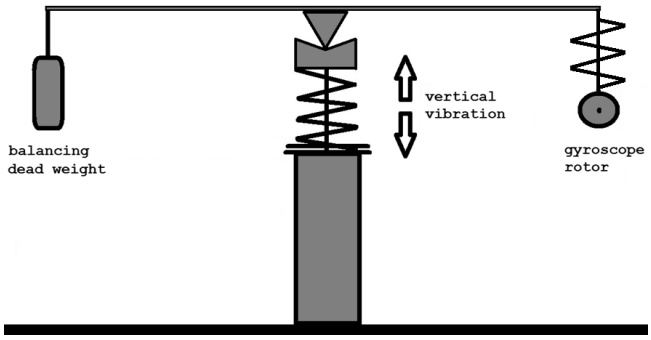

Fig. 1: Schematic of the compound pendulum developed by N.A. Kozyrev to measure a change in the weight of a gyroscope suspended from a balanced cross beam [6]. The preferred orientation of the cross beam appears to have been along the north/south direction, and that for the rotational axis of the gyroscope's rotor along the east/west direction. In some cases a weight change was detected by a small steady imbalance in the cross beam.

where c is the well-known speed of light in vacuum, e is the electronic charge in statcoulombs, and h is Plank's constant. The numerical value for the ratio e^2/h is 350 km/s. Kozyrev found by experiment that $c_2 \cong 700\,\text{km/s} = 2e^2/h = c/430 = \alpha c/\pi$.

A schematic for the compound pendulum developed by N.A. Kozyrev to measure c_2 is shown in Fig. 1 [6]. Kozyrev found that the weight of the gyroscope under certain conditions would change when there is a vertical vibration of the cross arm. Sometimes he observed a relative weight change on the order of 10^{-5}.

The objective of *this article* is to derive the effects of the neoclassical causal theory on a suspended gyroscope. We will find that the weight changes observed by N.A. Kozyrev may have been caused by the causal version of Newton's theory.

Parameter values and basis vectors

Numerical values for various parameters will be needed. Let m be the mass of the gyroscope's rotor, let R be its radius,

let ω_{rot} be its angular speed, let P_{rot} be the rotational period, let I_{rot} be the moment of inertia, let $\boldsymbol{J}_{rot}$ be the angular momentum vector, and let E_{rot} be the rotational energy. Typical numerical values for the parameters of an aircraft navigation gyroscope are [4]

$$\begin{aligned} &m = 0.1 \text{ kg},\\ &R = 2\times 10^{-2} \text{ m},\\ &\omega_{rot} = 2\pi 500 \text{ rad/s} = 3.14\times 10^{3} \text{ rad/s},\\ &P_{rot} = 2\pi/\omega_{rot} = 2\times 10^{-3} \text{ s},\\ &I_{rot} = mR^2 = 4\times 10^{-5} \text{ kg}\times\text{m}^2,\\ &J_{rot} = I_{rot}\omega_{rot} = 0.126 \text{ kg}\times\text{m}^2/s,\\ &E_{rot} = \frac{1}{2} I_{rot}\omega_{rot}^2 = 197 \text{ kg}\times m^2/s^2. \end{aligned} \tag{2}$$

Let the Earth be simulated by a spinning isotropic sphere of radius r_E, mass M_E, sidereal spin angular speed Ω_E, equatorial surface speed v_{eq}, moment of inertia I_E, surface gravitational scalar potential φ_E, surface gravitational field g_E, spin energy E_E, and spin angular momentum $\boldsymbol{J}_E$. Numerical values for the Earth's parameters are [1]

$$\begin{aligned} &G = 6.6732\times 10^{-11} \text{ N}\times\text{m}^2/\text{kg}^2,\\ &r_E = 6.37\times 10^{6} \text{ m},\\ &M_E = 5.98\times 10^{24} \text{ kg},\\ &\Omega_E = 7.29\times 10^{-5} \text{ rad/s},\\ &\mathrm{v}_{eq} = r_E\Omega_E = 4.65\times 10^{2} \text{ m/s},\\ &I_E = 8.02\times 10^{37} \text{ kg}\times\text{m}^2,\\ &\varphi_E = \frac{GM_E}{r_E} = 6.26\times 10^{7} \text{ m}^2/\text{s}^2,\\ &g_E = \frac{GM_E}{r_E^2} = 9.83 \text{ m/s}^2,\\ &E_E = \frac{1}{2} I_E\Omega_E^2 = 2.13\times 10^{29} \text{ kg}\times\text{m}^2/\text{s}^2,\\ &J_E = I_E\Omega_E = 5.85\times 10^{33} \text{ kg}\times\text{m}^2/\text{s}. \end{aligned} \tag{3}$$

Let (X, Y, Z) be the rectangular coordinates for an inertial frame-of-reference, let the Earth's center be at the origin, let the (X, Y) plane coincide with the equatorial plane, and let the axis of rotation coincide with the Z-axis. Let $\mathbf{e}_X$ be a unit vector directed outwardly along the X-axis, let $\mathbf{e}_Y$ be a unit vector directed outwardly along the Y-axis, and let $\mathbf{e}_Z$ be a unit vector directed outwardly along the Z-axis.

Let the spherical coordinates for an exterior field-point be (r,ϕ,λ), where r is the geocentric radial distance, ϕ is the azimuthal angle, and λ is the geocentric latitude. Let $\mathbf{e}_r$ be a unit vector directed upward along $\boldsymbol{r}$, let $\mathbf{e}_\phi$ be a unit vector directed towards the east, and let $\mathbf{e}_\lambda$ be a unit vector directed towards the north. The triad $(\mathbf{e}_r, \mathbf{e}_\phi, \mathbf{e}_\lambda)$ forms the basis for a right-handed system of orthogonal spherical coordinates.

Effects of a vertical vibration of a suspended gyroscope

Let the field-point be at the center of the rotor of an aircraft navigation gyroscope. Let λ be the geocentric latitude for the gyroscope. Let h be the rotor's height above the Earth's surface, let h_0 be a constant altitude, let h_1 be the vibration amplitude, and let ω_h be the angular speed for a vertical vibration. Then

$$h = h_0 + h_1\cos\omega_h t. \tag{4}$$

The time dependent geocentric radial distance becomes

$$r = r_E\left(1 + \frac{h_0}{r_E} + \frac{h_1}{r_E}\cos(\omega_h t)\right). \tag{5}$$

Let r_ϕ be the rotor's geocentric radius of gyration

$$r_\phi = r_E\cos\lambda\left(1 + \frac{h_0}{r_E} + \frac{h_1}{r_E}\cos(\omega_h t)\right). \tag{6}$$

Let $\mathbf{v}$ be the rotor's vector inertial velocity

$$\mathbf{v} = \mathbf{e}_r\mathrm{v}_r + \mathbf{e}_\phi\mathrm{v}_\phi + \mathbf{e}_\lambda\mathrm{v}_\lambda. \tag{7}$$

The formulas for v_r and v_ϕ are

$$\begin{aligned} &\mathrm{v}_r = \frac{dr}{dt} = -h_1\omega_h\sin(\omega_h t),\\ &\mathrm{v}_\phi = r_\phi\Omega_\phi = r_E\Omega_E\cos\lambda\left(1 + \frac{h_0}{r_E} + \frac{h_1}{r_E}\cos(\omega_h t)\right). \end{aligned} \tag{8}$$

Let E_r be the radial energy. If the radial energy is *constant*, then

$$\begin{aligned} constant = E_r &= \frac{1}{2} mv_r^2 - mg_E h =\\ &= \frac{1}{2} mh_1^2\omega_h^2\sin^2(\omega_h t) - mg_E(h_0 + h_1\cos(\omega_h t)). \end{aligned} \tag{9}$$

By using a trig identity for $\sin^2(\omega_h t)$, the time independent part of (9) becomes

$$constant = \frac{1}{4} mh_1^2\omega_h^2 - mg_E h_0. \tag{10}$$

Suppose a gyroscope is suspended by a spring of unstretched length ℓ_0 and spring constant k, as depicted in Fig. 2. Suppose the upper end of the spring is connected to a vibrator which can produce a time-dependent supporting force.

$$F_{up} = W + mh_{vib}\omega_{vib}^2\cos(\omega_{vib} t), \tag{11}$$

where W is the weight of the gyroscope. If the vibrator is turned off, $h_{vib} = 0$. In this case, the upper end of the spring is attached to a fixed solid point, and the system becomes a simple undriven harmonic oscillator.

Let $\delta\ell_0$ be the stretch of the spring when the gyroscope is attached. Then $k = W/\delta\ell_0 \cong mg_E/\delta\ell_0$, where g_E is the Earth's radial gravitational field at the surface. Let $\delta\ell_0 = h_0$. Then

$$k = \frac{mg_E}{h_0}. \tag{12}$$

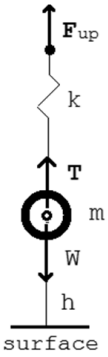

Fig. 2: Schematic for a forced harmonic oscillator; a rotor of mass m suspended by a spring of spring constant k with an upward supporting force F_{up}. Here T is the spring tension pulling up on the rotor, the weight W is the downward force of gravity on the rotor, and h is the height of the center of the rotor above the surface. Assume that the mass of the spring is negligible, and that the mass of the gyroscope approximately equals the mass of the rotor.

If the system is enclosed in a glass box, the damping of small amplitude free oscillations would be weak. The equation for an undamped harmonic oscillator is [7]

$$\frac{d^2h}{dt^2} + \omega_k^2 h = 0, \tag{13}$$

where

$$\omega_k^2 = \frac{k}{m} = \frac{g_E}{h_0}. \tag{14}$$

If $h_0 \cong 10^{-4}$ m, then $\omega_k \cong 313$ rad/s or 50 Hz. If $\omega_h = \omega_k$ and the constant of (10) is zero, the connection between h_1 and h_0 becomes

$$h_1 = 2h_0. \tag{15}$$

This shows that the constant h_0 is comparable with the amplitude h_1.

Now consider the forced harmonic oscillator. Suppose the vibrator is turned on and adjusted to an amplitude h_{vib} and angular speed ω_{vib}. In this case,

$$F_{up} = mg_E + mh_{vib}\omega_{vib}^2 \cos(\omega_{vib} t). \tag{16}$$

If $\omega_{vib} \cong \omega_k$, the system is at or near resonance [7]. At resonance, if the damping is small, the speed dh/dt is in phase with the driving force F_{up}, the average kinetic energy in the system is at a maximum, and the amplitude at the rotor h_1 can be many times greater than the driver amplitude h_{vib}.

The effects of vibration alone apply to any dead weight, because vibration alone does not depend on the rotation of the gyroscope's rotor. Gyroscopic forces do depend on the rotation of the rotor. Therefore, for a complete analysis, gyroscopic forces must be included.

Effects of gyroscopic forces

Gyroscopic forces cause precession and nutation [7, 8]. Precession is a steady revolution of the rotor around a vertical axis, and nutation is an up-down nodding motion of the rotor. The general problem for motions of a spinning rigid body can be quite complicated, but the problem is simplified for certain special cases. The case for "THE HEAVY SYMMETRICAL TOP WITH ONE POINT FIXED" is described in great detail by H. Goldstein [8, p. 213].

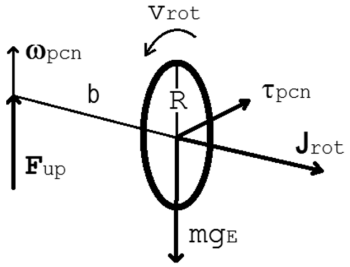

Fig. 3: Depiction of the gyroscopic forces acting on a rotor of mass m, radius R, and angular momentum vector $\mathbf{J}_{rot}$, which is supported by an upward force $\mathbf{F}_{up}$ at a distance b along the axel from the rotor's center to the support. Assume $\mathbf{J}_{rot}$ is in the horizontal plane. If $F_{up} = mg_E$, the precessional torque on the rotor $\tau_{pcn} = bmg_E$. In this case, the rotor precesses around the support with an angular speed $\omega_{pcn} = bmg_E/mR^2\omega_{rot}$.

Suppose the axel for a rotor is supported at a distance b from the center with an upward supporting force $\mathbf{F}_{up}$ and with the angular momentum vector $\mathbf{J}_{rot}$ released in the horizontal plane, as depicted in Fig. 3.

For a first case, suppose the supporting force is constant and equal to the weight, $F_{up} = mg_E$. Consider the case for slow precession without nutation.

Let ω_{pcn} be the precessional angular speed, and let v_{pcn} be the linear speed. Then the torque $\tau_{pcn} = bmg_E = J_{rot}\omega_{pcn}$. Solving for the angular speed gives $\omega_{pcn} = bg_E/R^2\omega_{rot}$.

If the distance $b = 0.1$ m, $R = 2\times10^{-2}$ m, and $\omega_{rot} = 3.14\times10^3$ rad/s, numerical values for ω_{pcn} and v_{pcn} are

$$\begin{aligned} \omega_{pcn} &= \frac{bg_E}{R^2\omega_{rot}} = 0.782 \quad \text{rad/s}, \\ \mathrm{v}_{pcn} &= b\omega_{pcn} = 7.82\times10^{-2} \quad \text{m/s}. \end{aligned} \tag{17}$$

Thus we find that the precessional speed for this case would be slow and constant at about 8 cm/s. Notice that this gyroscopic force supports the entire weight of the rotor.

Suppose the system is started with $\boldsymbol{J}_{rot}$ at a small initial angle $\delta\theta_0$ above the horizontal plane. Let h_{ntn} be the amplitude for nutation, which is the initial height above the horizontal plane. Then

$$h_{ntn} = b\tan\delta\theta_0. \tag{18}$$

When released, the rotor will precess with the angular speed ω_{pcn} of (17) and oscillate up and down with an upper maximum angle $\delta\theta_0$ and a lower minimum angle $\delta\theta_1$. Let ω_{ntn} be the angular speed for nutation. The formula for ω_{ntn} can be

found in [8, p. 221].

$$\omega_{ntn} = \frac{bg_E}{R^2\omega_{pcn}} = \omega_{rot} = 3.1\times10^3 \quad \text{rad/s}. \tag{19}$$

Thus we find that the frequency for nutation is the same as the frequency for the rotor, 500 Hz.

The formula for the difference $\sin\delta\theta_0 - \sin\delta\theta_1$ can be found in [7, p. 312].

$$\sin\delta\theta_0 - \sin\delta\theta_1 = \frac{2g_E b^3}{R^4\omega_{rot}^2}.$$

If $b = 0.1$ m, $R = 2\times10^{-2}$ m, and $\omega_{rot} = 3.14\times10^3$ rad/s, the numerical value for the difference becomes

$$\sin\delta\theta_0 - \sin\delta\theta_1 = 1.24\times10^{-2}, \tag{20}$$

the amplitude

$$h_{ntn} = 6.2\times10^{-4} \quad \text{m}, \tag{21}$$

and the linear speed for nutation becomes

$$\mathrm{v}_{ntn} = h_{ntn}\omega_{ntn}\sin(\omega_{ntn}t) \cong (1.9\ \text{m/s})\sin(\omega_{ntn}t). \tag{22}$$

Now let's change the length of the axel. Suppose the rotor's axel is extended on the other side of the support by the same distance b, and a dead weight that balances the cross beam is attached. If $\boldsymbol{J}_{rot}$ is directed outward from the supporting point, the dead weight would produce a torque equal in magnitude but opposite to the direction for τ_{pcn}, which would cancel the precessional motion. But such a balance would not cause any change in the nutational motion.

With the cross beam balanced in this manner, suppose the vibrator that supports the cross beam is turned on and adjusted to have an amplitude of h_{ntn} and an angular speed ω_{ntn}. This would induce an artificial nutation, but only if the gyroscope's rotor is spinning with an angular speed ω_{rot}. If the radial gravitational field contains a small time-dependent component with an angular speed near ω_{ntn}, there would be interesting interference effects and beat frequencies that could become visible in the balance of the cross beam.

The Earth's time-retarded transverse gravitational field

To satisfy the causality principle, the neoclassical causal theory postulates a new time-retarded transverse component for the Earth's gravitational field [1]. Let $\boldsymbol{g}_\phi$ be the Earth's time-retarded transverse component. The formula for the magnitude is [1]

$$g_\phi = C_\phi\left(1 - \frac{\Omega_\phi}{\Omega_E}\right)PS(r)\cos^2\lambda, \tag{23}$$

where the definition for the coefficient is

$$C_\phi = G\bar{\rho}r_E\frac{v_{eq}}{c_g}. \tag{24}$$

Here G is the gravity constant, r_E is the Earth's spherical radius, Ω_E is the Earth's sidereal angular speed, $\bar{\rho}$ is the Earth's mean mass density, c_g is the speed of propagation of the Earth's gravitational field, r is the geocentric radial distance to the field point, λ is the geocentric latitude for the field point, Ω_ϕ is the angular speed of the projection of the field point onto the equatorial plane, and $PS(r)$ is a power series representation for a triple integral over the Earth's volume.

The numerical value for C_ϕ with $c_g = c$ is

$$C_\phi = G\bar{\rho}r_E\frac{\mathrm{v}_{eq}}{c} = 3.635\times10^{-6} \quad \text{m/s}^2. \tag{25}$$

The formula for the power series is

$$PS(r) = \left(\frac{r_E}{r}\right)^3\left(C_0 + C_2\left(\frac{r_E}{r}\right)^2 + C_4\left(\frac{r_E}{r}\right)^4 + C_6\left(\frac{r_E}{r}\right)^6\right), \tag{26}$$

where the values for the coefficients are

$$\begin{aligned} C_0 &= 0.50889, \quad C_2 = 0.13931,\\ C_4 &= 0.01013, \quad C_6 = 0.14671. \end{aligned} \tag{27}$$

Let CPS_0 be the value for $PS(r_E)$. The definition and numerical value are

$$CPS_0 = C_0 + C_2 + C_4 + C_6 = 0.805. \tag{28}$$

Let $\boldsymbol{J}_Z$ be the geocentric angular momentum for the rotor, defined as

$$J_Z = mr_\phi^2\Omega_\phi \tag{29}$$

By conservation of angular momentum,

$$constant = \frac{J_z}{m} = r_\phi^2\Omega_\phi = r_E^2\Omega_E\cos^2\lambda \tag{30}$$

Solving (30) for Ω_ϕ gives

$$\Omega_\phi \cong \Omega_E\left(1 - 2\frac{h_0}{r_E} - 2\frac{h_1}{r_E}\cos(\omega_h t)\right) \tag{31}$$

Then the difference

$$1 - \frac{\Omega_\phi}{\Omega_E} = 2\frac{h_0}{r_E} + 2\frac{h_1}{r_E}\cos(\omega_h t). \tag{32}$$

Substituting (32) into (23) produces

$$g_\phi = C_\phi\left(2\frac{h_0}{r_E} + 2\frac{h_1}{r_E}\cos(\omega_h t)\right)PS(r)\cos^2\lambda. \tag{33}$$

The numerical value for g_ϕ with $c_g = c$, $r = r_E$, $h_0 = h_1 = 10^{-4}$ m, and $\lambda = 60^\circ$, is

$$g_\phi = \left(2.3\times10^{-17}\ \text{m/s}^2\right)\left(1 + \cos(\omega_h t)\right). \tag{34}$$

This result shows that the time-retarded transverse gravitational field for a suspended gyroscope is totally negligible.

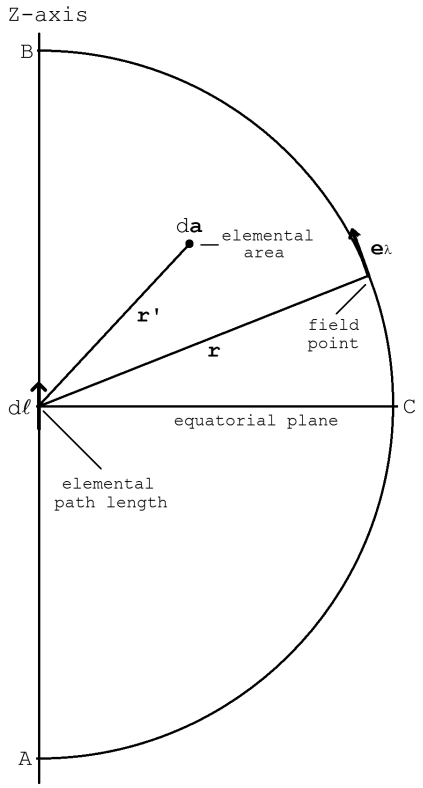

Fig. 4: Depiction of the semicircular area to be used for Stoke's theorem. The contour for the line integral is $A \to B \to C \to A$. Here $d\boldsymbol{\ell}$ is an elemental path length vector, d**a** is an elemental area vector, and $\mathbf{e}_\lambda$ is a unit vector for λ. The field-point is at **r** and the elemental area d**a** is at **r**′.

The Earth's time-retarded transverse vector potential

Let $\boldsymbol{A}$ be the vector potential for $\boldsymbol{g}_\phi$. Then by definition

$$\boldsymbol{g}_\phi = \boldsymbol{\nabla} \times \boldsymbol{A}. \tag{35}$$

Units for A are m^2/s^2, the same as the units for the scalar potential. Because the divergence of $\boldsymbol{g}_\phi$ is zero, the divergence of $\boldsymbol{A}$ must also be zero, which means that $\boldsymbol{A}$ cannot have a component directed along $\boldsymbol{e}_r$. Consequently, $\boldsymbol{A}$ must be directed along $\boldsymbol{e}_\lambda$.

The needed elemental vectors $d\boldsymbol{\ell}$ and d**a** for integration using Stoke's theorem are depicted in Fig. 4. Stoke's theorem states that the line integral of $\boldsymbol{A} \bullet d\boldsymbol{\ell}$ around a closed contour equals the surface integral of $\boldsymbol{\nabla} \times \boldsymbol{A} \bullet d\mathbf{a}$ over the surface bounded by the contour. It is symbolically written as

$$\oint \boldsymbol{A} \bullet \mathrm{d}\boldsymbol{\ell} = \iint \boldsymbol{\nabla} \times \boldsymbol{A} \bullet \mathrm{d}\mathbf{a}. \tag{36}$$

Consider the closed contour depicted in Fig. 4: $A \to B \to C \to A$. The left side of (36) becomes

$$\oint_{A \to B} \boldsymbol{A} \bullet d\boldsymbol{\ell} = 0, \qquad \oint_{B \to C \to A} \boldsymbol{A} \bullet d\boldsymbol{\ell} = A_\lambda \pi r. \tag{37}$$

The right side of (36) becomes

$$\iint \boldsymbol{\nabla} \times \boldsymbol{A} \bullet \mathrm{d}\mathbf{a} = g_\phi \iint r' dr' d\lambda' = g_\phi \frac{\pi}{2} r^2. \tag{38}$$

Next comes the solution

$$A_\lambda = \frac{1}{2} r g_\phi = A_0 \cos^2 \lambda \, PS'(r) \left(2\frac{h_0}{r_E} + 2\frac{h_1}{r_E} \cos(\omega_h t) \right) \tag{39}$$

where the definition for A_0 and its numerical value with $c_g = c$ and the definition for the power series for A_λ are

$$\begin{aligned} A_0 &= \frac{C_\phi r_E}{2} = 11.6 \quad \mathrm{m^2/s^2}, \\ PS'(r) &= \frac{r}{r_E} PS(r) = \\ &= C_0 \left(\frac{r_E}{r}\right)^2 + C_2 \left(\frac{r_E}{r}\right)^4 + C_4 \left(\frac{r_E}{r}\right)^6 + C_6 \left(\frac{r_E}{r}\right)^8. \end{aligned} \tag{40}$$

The formula that connects $\boldsymbol{g}_\lambda$ to the time-dependence of $\boldsymbol{A}_\lambda$ is [9, p. 219].

$$\begin{aligned} g_\lambda &= -\frac{1}{\mathrm{v}_k} \frac{dA_\lambda}{dt} = \\ &= 2\frac{A_0}{\mathrm{v}_k} \cos^2 \lambda \left(\frac{h_1 \omega_h}{r_E} PS'(r_E) \sin(\omega_h t) - \right. \\ &\quad \left. - \frac{dPS'}{dt} \left(\frac{h_0}{r_E} + \frac{h_1}{r_E} \cos(\omega_h t) \right) \right), \end{aligned} \tag{41}$$

where v_k is the "induction speed" for the neoclassical causal theory.

The numerical value for the average induction speed has been found to be [1]

$$\overline{\mathrm{v}}_k \cong 5 \times 10^3 \ \mathrm{m/s}. \tag{42}$$

The coefficient A_0 is inversely proportional to c_g. It is interesting to notice that A_0/v_k with $c_g = c$ is inversely proportional to $c\mathrm{v}_k$, and that

$$\sqrt{c\mathrm{v}_k} \cong 11 \times 10^5 \ \mathrm{m/s} = 1.7 c_2, \tag{43}$$

where c_2 is Kozyrev's secondary universal speed, the one that is to be associated with rotational motion [4].

Let CPS'_0 be the value for PS' at $r = r_E$.

$$CPS'_0 = PS'(r_E) = C_0 + C_2 + C_4 + C_6 = 0.805. \tag{44}$$

The value for dPS'/dt evaluated at $r = r_E$ is

$$\begin{aligned} \left.\frac{dPS'}{dt}\right|_{r=r_E} &= (2C_0 + 4C_2 + 6C_4 + 8C_6) \frac{h_1 \omega_h}{r_E} \sin(\omega_h t) = \\ &= 2.81 \frac{h_1 \omega_h}{r_E} \sin(\omega_h t). \end{aligned} \tag{45}$$

The formula for g_λ to first order in h_1/r_E reduces to

$$g_\lambda \cong C_\lambda \cos^2 \lambda \frac{h_1 \omega_h}{\mathrm{v}_k} \sin(\omega_h t), \tag{46}$$

where the definition and numerical value with $c_g = c$ for C_λ are

$$C_\lambda = 0.805 \times 2\frac{A_0}{r_E} = 0.805 C_\phi = 2.926 \times 10^{-6} \ \mathrm{m/s^2}, \quad (47)$$

and C_ϕ is given by (24).

If $h_1 = 10^{-4}$ m, $\omega_h = \omega_{rot}$, $\mathrm{v}_k = 5$ km/s, and $\lambda = 60^\circ$, the numerical value for g_λ reduces to

$$g_\lambda = (4.6 \times 10^{-11} \ \mathrm{m/s^2}) \sin(\omega_h t) \quad (48)$$

This result shows that the vector potential can produce a relatively large value for the north/south transverse gravitational field. The ratio for g_λ / g_ϕ, with g_ϕ from (34), is on the order of

$$\frac{g_\lambda}{g_\phi} \sim 2 \times 10^6. \quad (49)$$

Secondary radial induction field

The analogous Lorentz force law for gravity [1, 2] states that a north/south transverse gravitational field can induce a radial gravitational field. Let $\boldsymbol{g}_{ind}$ be the induced gravitational field. Then

$$\boldsymbol{g}_{ind} = \frac{\mathbf{v}}{\mathrm{v}_k} \times \boldsymbol{g} = \frac{1}{\mathrm{v}_k} \begin{vmatrix} \mathbf{e}_r & \mathbf{e}_\phi & \mathbf{e}_\lambda \\ \mathrm{v}_r & \mathrm{v}_\phi & \mathrm{v}_\lambda \\ g_r & g_\phi & g_\lambda \end{vmatrix}. \quad (50)$$

The induced gravitational field along $\mathbf{e}_r$ is the only one of the components that can change the weight of the rotor.

$$\mathbf{e}_r g_{ind} = \mathbf{e}_r \left(\frac{\mathrm{v}_\phi}{\mathrm{v}_k} g_\lambda - \frac{\mathrm{v}_\lambda}{\mathrm{v}_k} g_\phi \right) \cong \mathbf{e}_r \frac{\mathrm{v}_\phi}{\mathrm{v}_k} g_\lambda. \quad (51)$$

Substituting (8) and (42) into (51) gives

$$g_{ind} \cong C_{ind} \sin(\omega_h t), \quad (52)$$

where

$$C_{ind} = C_\lambda \frac{h_1 \omega_h}{\mathrm{v}_k} \frac{\mathrm{v}_{eq}}{\mathrm{v}_k} \cos^3 \lambda \quad (53)$$

If $\lambda = 60^\circ$, $h_1 = 10^{-4}$ m, $\omega_h = \omega_{rot}$, and $\mathrm{v}_k = 5$ km/s, the numerical value for C_{ind} reduces to

$$C_{ind} = 2.1 \times 10^{-12} \quad \mathrm{m/s^2}. \quad (54)$$

This result predicts a very small value for g_{ind}, but it is close to the order of magnitude for g_λ, which is predicted to be about 10^6 times g_ϕ. There may be some hidden effect that enhances g_{ind} by 10^6, in particular the nutation effects of (21) and (22). This question can be resolved only by experiment.

Conclusions and recommendations

It seems plausible but not proven that the weight changes observed by N.A. Kozyrev may have been caused by the neoclassical causal theory. Modern experimental techniques using digital electronics, sensitive strain gauges, sensitive accelerometers, and computer controls, can greatly increase the sensitivity and reliability of laboratory instruments. If an instrument that can detect the Earth's time-retarded transverse vector potential is developed with a precision of about 1 part in 1000, the theory can be worked backwards to provide a measured value for the speed of the Earth's gravitational field. To accomplish this end, a dedicated effort to develop an instrument, and comprehensive systematic studies using such an instrument, are highly recommended.

Acknowledgements

I thank Patrick L. Ivers for reviewing the manuscript for this article and suggesting improvements. I also thank Dr. Dmitri Rabounski for providing the schematic drawing for N. A. Kozyrev's compound pendulum.

Submitted on August 18, 2013 / Accepted on August 25, 2013
Updated with corrections on October 10, 2013

References

1. Hafele J.C. Earth Flyby Anomalies Explained by a Time-Retarded Causal Version of Newtonian Gravitational Theory. *The Abraham Zelmanov Journal*, 2012, v.5, 134–187.
2. Hafele J.C. Causal Version of Newtonian Theory by Time-Retardation of the Gravitational Field Explains the Flyby Anomalies. *Progress in Physics*, 2013, v.3, 3–8.
3. Hafele J.C. Comment on N.A. Kozyrev's "Possibility of Experimental Study of the Properties of Time". *Letters to Progress in Physics*, 2013, v.3, L1.
4. Kozyrev N.A. Possibility of Experimental Study of the Properties of Time. *The Abraham Zelmanov Journal*, 2012, v.5, 188–220.
5. Fine-structure constant. Wikipedia, http://en.wikipedia.org/wiki/Fine-structure_constant
6. Private communication from Dr. Dmitri Rabounski.
7. Becker R.A. Introduction to Theoretical Mechanics. McGraw-Hill, New York, 1954.
8. Goldstein H. Classical Mechanics. Addison-Wesley, Reading, Massachusetts, 1980.
9. Jackson J.D. Classical Electrodynamics. Wiley, New York, 1975.

Non-Linear Effects in Flow in Porous Duct

Hafeez Y. Hafeez*, Jean Bio-Chabi Orou† and Omololu A. Ojo‡

*Physics Department, Federal University Dutse, Nigeria. E-mail: hafeezyusufhafeez@yahoo.com
†University Abomey-Calavi, Republic of Benin. E-mail: Jchabi@yahoo.fr
‡African University of Science and Technology, Abuja – Nigeria. E-mail: prayerz@yahoo.com

In general, it is assumed in some non viscous flows that the flow velocity is constant at a cross-section. In this paper, we impose more realistic boundary conditions by, for example, introducing viscosity, and suction at walls, the net mass flow will change since the continuity equation must hold. The convective acceleration terms will be products of variables such that a non-linear behaviour will take place in the flow. The work will consist of deriving all the equations and parameters needed to described this kind of flow. An approximate analytic solution for the case of small Reynold number Re is discussed using perturbation techniques. Expression for the velocity components and pressure are obtained. The governing non-linear differential equation that cannot be solved analytically is solved numerically using Runge-Kutta Program and the graphs of axial and lateral velocity profiles are drawn.

1 Introduction

The problems of fluid flow through porous duct have arouse the interest of Engineers and Mathematicians, the problems have been studied for their possible applications in cases of membrane filtration, transpiration cooling, gaseous diffusions and drinking water treatment as well as biomedical engineering. Such flows are very sensitive to the Reynold number.

Berman was the first researcher who studied the problem of steady flow of an incompressible viscous fluid through a porous channel with rectangular cross section, when the Reynold number is low and the perturbation solution assuming normal wall velocity to be equal was obtained [1].

Sellars [2], extended the problem studied by Berman by using very high Reynold numbers.

Also wall suctions were recognize to stabilize the boundary layer and critical Reynold number for natural transition 46130 was obtained [3]. The stabilization effects of wall suction is due to the change of mean velocity profiles.

In the review of Joslin [4], it is also noticed that the uniform wall suction is not only a tool for laminar flow control but can also be used to damped out already existing turbulence.

The effects of Hall current on the steady Hartman flow subjected to a uniform suction and injection at the boundary plates has been studied [5].

Other reviews of flow in porous duct tend to focus only on one specific aspect of the subject at a time such as membrane filteration [8], the description of boundary conditions [6] and the existence of exact solutions [7].

In this paper, we consider the steady two-dimensional laminar flow of an incompressible viscous fluid between two parallel porous plates with equal suction and assume that the wall velocity is non uniform.

2 Formulation of the problem

The steady laminar flow of an incompressible viscous fluid between two parallel porous plates with an equal suction at walls and non uniform cross flow velocity is considered. The well known governing equations of the flow are:

Continuity equation

$$\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}=0. \tag{1}$$

Momentum equations (without body force)

$$u\frac{\partial u}{\partial x}+v\frac{\partial u}{\partial y}=-\frac{1}{\rho}\frac{\partial p}{\partial x}+\nu\left(\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}\right), \tag{2}$$

$$u\frac{\partial v}{\partial x}+v\frac{\partial v}{\partial y}=-\frac{1}{\rho}\frac{\partial p}{\partial y}+\nu\left(\frac{\partial^2 v}{\partial x^2}+\frac{\partial^2 v}{\partial y^2}\right). \tag{3}$$

Let us consider channel flow between uniformly parallel plates with equal suction. Assuming that we are far downstream of the entrance, the boundary conditions can be defined as

$$y=h, u=0, v=v_w, \tag{4}$$

$$y=-h, u=0, v=-v_w. \tag{5}$$

Let $\overline{u}(0)$ denote the average axial velocity at an initial section ($x=0$). Then it is clear from a gross mass balance that $\overline{u}(x)$ will differ from $\overline{u}(0)$ by the amount $\frac{v_w}{h}x$. This observation led Berman(1953) to formulate the following relation for the stream in the channel [9].

$$\psi(x,y)=(h\overline{u}(0)-v_w x)\,f(y^*). \tag{6}$$

Where $y^*=\frac{y}{h}$, $\psi(x,y)$ is a stream function,$\overline{u}(0)$ is initial average axial velocity and f is dimensionless function to be determined. The velocity components follow immediately from

the definition of ψ:

$$u(x,y^*)=\frac{\partial\psi}{\partial y}=\left(\overline{u}(0)-\frac{v_w x}{h}\right)f'(y^*)=\overline{u}(x)f'(y^*), \quad (7)$$

$$v(x,y^*)=-\frac{\partial\psi}{\partial x}=v_w f(y^*)=v(y). \quad (8)$$

The stream function must now be made to satisfy the momentum equations (2) and (3) for steady flow (2) and (3) will now become

$$u\frac{\partial u}{\partial x}+\frac{v}{h}\frac{\partial u}{\partial y^*}=-\frac{1}{\rho}\frac{\partial p}{\partial x}+\nu\left(\frac{\partial^2 u}{\partial x^2}+\frac{1}{h^2}\frac{\partial^2 u}{\partial y^{*2}}\right), \quad (9)$$

$$u\frac{\partial v}{\partial x}+\frac{v}{h}\frac{\partial v}{\partial y^*}=-\frac{1}{\rho h}\frac{\partial p}{\partial y^*}+\nu\left(\frac{\partial^2 v}{\partial x^2}+\frac{1}{h^2}\frac{\partial^2 v}{\partial y^{*2}}\right). \quad (10)$$

Using (7) and (8) in (9) and (10), the momentum equations reduces to,

$$-\frac{1}{\rho}\frac{\partial p}{\partial x}=\left(\overline{u}(0)-\frac{v_w x}{h}\right)\left(\frac{v_w}{h}\left(ff''-f'^2\right)-\frac{\nu}{h^2}f'''\right), \quad (11)$$

$$-\frac{1}{\rho h}\frac{\partial p}{\partial y^*}=\frac{v_w^2}{h}ff'-\frac{\nu v_w}{h^2}f''. \quad (12)$$

Now differentiating (12) w.r.t x, we get

$$\frac{\partial^2 p}{\partial x\partial y^*}=\frac{\partial^2 p}{\partial x\partial y}=0. \quad (13)$$

Differentiating (11) w.r.t y^*, we get

$$\frac{\partial^2 p}{\partial x\partial y^*}=\left(\overline{u}(0)-\frac{v_w x}{h}\right)\frac{d}{dy^*}\left(\frac{v_w}{h}\left(ff''-f'^2\right)-\frac{\nu}{h^2}f'''\right). \quad (14)$$

From (13), (14) can be written as

$$\frac{d}{dy^*}\left(\frac{v_w}{h}\left(ff''-f'^2\right)-\frac{\nu}{h^2}f'''\right)=0, \quad (15)$$

$$\frac{v_w}{h}\left(ff'''-f'f''\right)-\frac{\nu}{h^2}f''''=0.$$

Let the suction Reynold number be $Re=\frac{hv_w}{\nu}$ and substitute into above expression, we get

$$f''''+Re\left(f'f''-ff'''\right)=0. \quad (16)$$

(16) has no known analytic-closed form solution, but it can be integrated once i.e integrate (16) w.r.t y^*, we get

$$f'''+Re\left(f'^2-ff''\right)=K=const. \quad (17)$$

The boundary conditions on $f(y^*)$ 0f (4) and (5) can now be written as,

$$f(1)=1, f(-1)=-1, f'(1)=0, f'(-1)=0. \quad (18)$$

Hence, the solution of the equations of motion and continuity is given by non-linear fourth order differential equation (16) subject to the boundary condition (18).

3 Results

3.1 Approximate analytic solution (perturbation)

The non-linear ordinary differential equation (16) subject to condition (18) must in general be integrated numerically. However for special case when "Re" is small, approximate analytic results can be obtained by the use of a regular perturbation approach. Note that perturbation method has been used because the equations (16 and 18) are non-linear by using that technique, we get a linear approximated version of the true equations. The solution of $f(y^*)$ may be expanded in power of Re [10]

$$f(y^*)=\sum_{n=0}^{\infty}Re^n f_n(y^*) \quad (19)$$

where $f_n(y^*)$ satisfies the symmetric boundary conditions

$$f_0(0)=f_0'(1)=f_0''(0)=0, \quad f_0(1)=1 \quad (20)$$

and

$$f_n(0)=f_n'(1)=f_n''(0)=0, \quad f_n(1)=1. \quad (21)$$

Here f_n are independent of Re. Substituting (19) in (16), we get

$$\left(f_0''''+Ref_1''''+Re^2f_2''''\right)+Re\Big[\left(f_0'+Ref_1'+Re^2f_2'\right)\left(f_0''+Ref_1''+Re^2f_2''\right)-\left(f_0+Ref_1+Re^2f_2\right)\left(f_0'''+Ref_1'''+Re^2f_2'''\right)\Big]=0.$$

Equating coefficients of Re, we get

$$f_o''''=0, \quad (22)$$

$$f_1''''+f_o'f_o''-f_of_o'''=0, \quad (23)$$

$$f_2''''+f_o'f_1''+f_1'f_o''-f_of_1'''-f_1f_o'''=0. \quad (24)$$

The solution of (22) is of the form

$$f_o(y^*)=\frac{Ay^{*3}}{6}+\frac{By^{*2}}{2}+Cy^*+D,$$

where A,B,C and D are constants.

Applying the boundary condition (20) to the above equation, we get

$$f_o(y^*)=\frac{1}{2}\left(3y^*-y^{*3}\right). \quad (25)$$

The solutions of Eq (23) and (24) subject to the boundary condition (21), are:

$$f_1(y^*)=-\frac{1}{280}\left(y^{*7}-3y^{*3}-2y^*\right), \quad (26)$$

$$f_2(y^*)=\frac{1}{1293600}\times\left(14y^{*11}-385y^{*9}+198y^{*7}+876y^{*3}-703y^*\right). \quad (27)$$

Hence, the first order perturbation solution for $f(y^*)$ is

$$f'(y^*) = f_o(y^*) + Ref_1(y^*),$$
$$f^1(y^*) = \frac{1}{2}\left(3y^* - y^{*3}\right) - \frac{Re}{280}\left(y^{*7} - 3y^{*3} - 2y^*\right). \quad (28)$$

The second order perturbation of solution for $f(y^*)$ is

$$f^2(y^*) = f_o(y^*) + Ref_1(y^*) + Re^2 f_2(y^*),$$

$$f^2(y^*) = \frac{1}{2}\left(3y^* - y^{*3}\right) - \frac{Re}{280}\left(y^{*7} - 3y^{*3} - 2y^*\right) + \frac{Re^2}{1293600}\left(14y^{*11} - 385y^{*9} + 198y^{*7} + 876y^{*3} - 703y^*\right). \quad (29)$$

Hence, the first order expression for the velocity components are:

$$u(x, y^*) = \left[\overline{u}(0) - \frac{v_w x}{h}\right] f'(y^*) = \left[\overline{u}(0) - \frac{v_w x}{h}\right]\frac{3}{2}\left(1 - y^{*2}\right)\left(1 - \frac{Re}{420}\left(2 - 7y^{*2} - 7y^{*4}\right)\right), \quad (30)$$

$$v(x, y^*) = v_w f(y^*) = v_w\left[\frac{1}{2}\left(3y^* - y^{*3}\right) - \frac{Re}{280}\left(y^{*7} - 3y^{*3} - 2y^*\right)\right]. \quad (31)$$

For pressure distribution, from Eq. (11) we get

$$\frac{h^2}{\rho\nu}\frac{\partial p}{\partial x} = \left[\overline{u}(0) - \frac{v_w x}{h}\right]\left[f'''(y^*) + Re\left(f'^2(y^*) - f(y^*)f''(y^*)\right)\right],$$

and since $f'''(y^*) + Re\left(f'^2(y^*) - f(y^*)\right)f''(y^*)) = K$, from (17), we have:

$$\frac{\partial p}{\partial x} = \frac{K\rho\nu}{h^2}\left[\overline{u}(0) - \frac{v_w x}{h}\right] = \frac{K\mu}{h^2}\left[\overline{u}(0) - \frac{v_w x}{h}\right]. \quad (32)$$

Now, from Eq. (12), we have

$$\frac{\partial p}{\partial y^*} = \frac{\mu v_w}{h} f''(y^*) - \rho\nu^2 f(y^*)f'(y^*). \quad (33)$$

Since $dp = \frac{\partial p}{\partial x}dx + \frac{\partial p}{\partial y^*}dy^*$, then

$$dp = \frac{K\mu}{h^2}\left[\overline{u}(0) - \frac{v_w x}{h}\right]dx + \left[\frac{\mu v_w}{h} f''(y^*) - \rho\nu^2 f(y^*)f'(y^*)\right]dy^*. \quad (34)$$

Integrating (34), we get

$$p(x, y^*) = p(0, 0) - \frac{\rho\nu^2}{2} f^2(y^*) + \frac{K\mu}{h^2}\left[\overline{u}(0)x - \frac{v_w x^2}{2h}\right] + \frac{\mu v_w}{h}\left[f'(y^*) - f'(0)\right]. \quad (35)$$

The pressure drop in the major flow direction is given by

$$p(x, 0) - p(x, y^*) = \frac{K\mu}{h^2}\left[\frac{v_w x^2}{2h} - \overline{u}(0)x\right]. \quad (36)$$

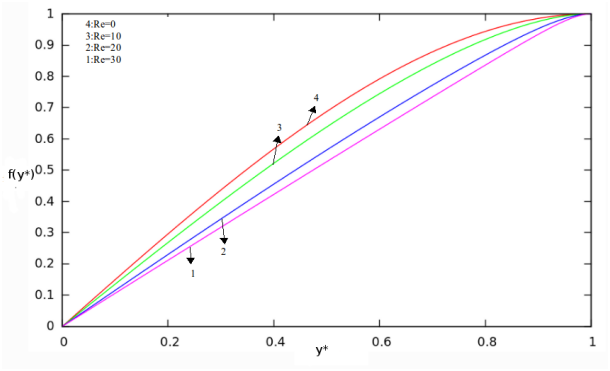

Fig. 1: Lateral velocity profiles for flow between parallel plates with equal suctions for different values of Re.

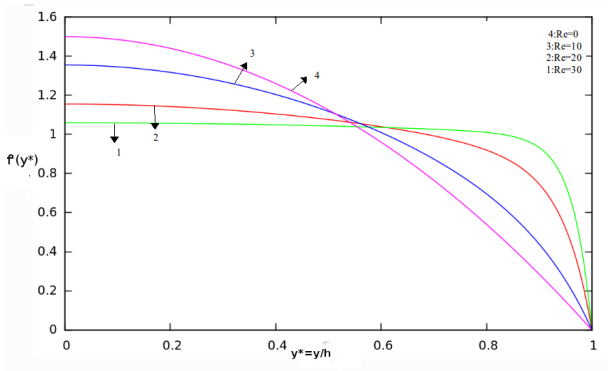

Fig. 2: Axial velocity profiles for flow between parallel plates with equal suctions for different values of Re.

3.2 Numerical solution

The approximate results of the previous section are not reliable when the Reynold number is not small. To obtain the detail information on the nature of the flow for different values of Reynold number (i.e. $Re = 0, 10, 20, 30$), a numerical solution to the governing equations is necessary. The Runge-Kutta program App.C is used to solve Eq. (17) numerically. One initial condition and constant (K) are unknown; i.e. starting at $y^* = 1$, then $f''(1)$ and K were guessed and the solution double-iterated until $f(-1) = -1$ and $f'(-1) = 0$. The most complete sets of profiles are shown in the figs. 1 and 2.

4 Discussion

The velocity profiles have been drawn for different values of Reynold number (i.e. $Re = 0, 10, 20, 30$). The shapes change smoothly with Reynold number and show no odd or unstable behaviour. Suction tends to draw the profiles toward the wall. From fig. (1), it is observed that for $Re > 0$ in the region $0 \le y^* \le 1$, $f(y^*)$ decreases with the increase of Reynold number Re. Also from fig. (2), it is observed that, for $Re > 0$,

then $f'(y^*)$ decreases with an increase of the Reynold number in the range of $0 \leq y^* \leq 1$.

5 Conclusion

In this paper, a class of solutions of laminar flow through porous duct has been presented. Numerical approach is necessary for arbitrary values of Re. Also, when a cross flow velocity along the boundary is not uniform, a numerical technique is necessary to solve Eq. (2) and (3). Also, from the results obtained in this article, we can now conclude that, the non-linear effects of a flow of the porous duct is due to non uniform cross flow velocity and non vanishing terms of convective acceleration of momentum equations. The perturbation solution obtained for this problem reduces to the results of Berman [1].

Acknowledgement

The authors wish to thank Dr. Chifu E. Ndikilar for his time, useful discussions and help in preparing the numerical solution to this problem.

Nomenclature

A,B,C,D: Constants
K: Arbitrary Constant
f: Dimensionless function representing lateral velocity profile
h: Height of the channel
P: Pressure
x: Axial distance
y: Lateral distance
v_w: Lateral wall velocity
u(x,y): Axial velocity component
v(x,y): Lateral velocity component
$y^* = \frac{y}{h}$: Dimensionless lateral distance
$Re = \frac{v_w h}{\nu}$: Wall Reynold number

Greek Symbols

μ: Shear viscosity
ν: Kinematic viscosity
ρ: Fluid density
$\psi(x, y)$: Stream function.

Submitted on: August 23, 2013 / Accepted on: September 03, 2013

References

1. Berman A.S. Laminar flow in channels with porous walls. *Journal of Applied Physics* 1953, v. 24, 1232–1235.
2. Sellars J.R. Laminar flow in channels with porous walls at high suction Reynold number. *Journal of Applied Physics* 1955, v. 26, 489.
3. Drazin P.G. and Reid W.H. Hydrodynamic stability, Cambridge University Press, (1981).
4. Joslin R.D. Aircraft laminar flow control. *Annual Review of Fluid Mechanics* 1998, v. 30, 1–29.
5. Abu-El-Dhat E. Hartman flow with uniform suction and injection at the boundary plate. M.Sc. Thesis, Helwan University, Egypt, (1993).
6. Sahraoui M. and Kaviany M. Slip and no-slip velocity boundary-conditions at interface of porous, plain media. *International Journal of Heat and Mass Transfer*, 1992, v. 35(4), 927–943.
7. Wang C.Y. Exact solutions of the steady-state Navier-Stokes equations. *Annual Review of Fluid Mechanics*, 1991, v. 23, 159–178.
8. Belfort G. Fluid-mechanics in membrane filtration – recent developments. *Journal fo Membrane Science*, 1989, v. 40, 123–147.
9. Frank White. Viscous fluid flow, 2nd edition, McGraw-Hill company New York, (1991).
10. Ganesh S. and Krishnambal S. Magnetohydrodynamics flow of viscous fluid between two parallel plates. *Journal of Applied Science* 2006, v. 6(11), 2420–2425.

Understanding the Dirac Equation and the Electron-Vacuum System

William C. Daywitt

National Institute for Standards and Technology (retired), Boulder, Colorado. E-mail: wcdaywitt@me.com

It has been close to a century since the Dirac equation first appeared, but it has yet to be understood on an intuitive, fundamental level. The reason for this lack of understanding is twofold: the equation is expressed in terms of the *secondary* constant $\hbar$; and the vacuum state and its coupling to the electron particle have not been developed as part of the electron model. What follows briefly reviews the vacuum coupling and illustrates it by deriving the Schrödinger and Pauli equations as derivatives of the Dirac equation, and by explaining the zitterbewegung response that is a vacuum dynamic associated with the coupling force. It is argued that the fields of quantum electrodynamics have as their origin the degenerate vacuum state.

1 Introduction

The Dirac electron defined here is a massive "point" charge $(-e_*, m)$ that obeys the Dirac equation and is coupled to the negative-energy Planck vacuum (PV) continuum via the two-term coupling force [1]

$$\frac{e_*^2}{r^2} - \frac{mc^2}{r} \tag{1}$$

the massive charge exerts on the PV. The electron Compton radius $r_c\,(= e_*^2/mc^2)$ is that radius from the center of the massive charge (in its rest frame) to the radius r_c where the coupling force vanishes. The bare charge $(-e_*)$ itself is massless, while the electron mass m results from the bare charge being driven by the zero-point electromagnetic field [2] [3]; corresponding to which is a vanishingly small sphere containing the driven charge whose center defines the center of both the driven charge and its derived mass. It is from the center of this small sphere that the position operator $\mathbf{r}$ for the massive charge and the electron-vacuum complex is defined and from which the radius r in (1) emerges.

The PV model of the complete electron consists of two interdependent dynamics, the dynamics of the massive charge in the previous paragraph and the dynamics of the PV continuum to which the massive charge is coupled. An example of the latter dynamic is the (properly interpreted) zitterbewegung [4] [1] that represents a harmonic-oscillator-type excitation taking place at the $r = r_c$ sphere surrounding the massive point charge, an oscillation resulting from the vacuum response to the vanishing of (1) at r_c. The point-like nature of the massive charge, in conjunction with the continuum nature of the PV, are what give the electron its so-called wave-particle-duality. Mathematically, the electron's wave nature is apparent from the fact that the spinor solutions to the Dirac equation are spinor *fields*, and it is upon these fields that the covariant gradient operator

$$\partial_\mu = \frac{\partial}{\partial x^\mu} = \left(\frac{\partial}{c\partial t}, \nabla\right) \tag{2}$$

operates. Thus the spinors are associated with PV distortion — with no distortion the gradients vanish, resulting in null spinors and the dissolution of (3).

The free-particle Dirac equation can be expressed in the form (from (A10) in Appendix A)

$$ir_c \frac{\partial}{c\partial t}\begin{pmatrix}\phi\\ \chi\end{pmatrix} + \begin{pmatrix}\vec{\sigma}\cdot ir_c\nabla\chi\\ \vec{\sigma}\cdot ir_c\nabla\phi\end{pmatrix} = \begin{pmatrix}\phi\\ -\chi\end{pmatrix} \tag{3}$$

in terms of the single constant r_c, a constant that normalizes the operator in (2). The free-space particle solution ϕ, and the negative-energy vacuum solution χ, for this electron-vacuum system are 2×1 spinors and $\vec{\sigma}$ is the Pauli 2×2 vector matrix. The spinor solutions from the two simultaneous equations in (3) are strongly coupled by the inverted χ-ϕ spinor configuration of the second term, showing the vacuum state to be an integral part of the electron phenomenon. (It will be seen that this coupling is even present in the nonrelativistic Schrödinger equation.) The negative spinor $(-\chi)$ on the right is a manifestation of the negative-energy nature of the vacuum. Equation (3) expresses the Dirac equation in terms of the normalized PV gradients on the left of the equal sign.

What follows illustrates the previous ideas by reiterating the standard development of the free-particle Schrödinger equation and the minimal coupling substitution leading to the Pauli equation.

2 Schrödinger equation

The Dirac-to-Schrödinger reduction [5, p. 79] begins with eliminating the high-frequency components from (3) by assuming

$$\begin{pmatrix}\phi\\ \chi\end{pmatrix} = \begin{pmatrix}\phi_0\\ \chi_0\end{pmatrix} e^{-imc^2t/\hbar} = \begin{pmatrix}\phi_0\\ \chi_0\end{pmatrix} e^{-ict/r_c} \tag{4}$$

where ϕ_0 and χ_0 are slowly varying functions of time compared to the exponentials. Inserting (4) into (3) gives

$$ir_c \frac{\partial}{c\partial t}\begin{pmatrix}\phi_0\\ \chi_0\end{pmatrix} + \begin{pmatrix}\vec{\sigma}\cdot ir_c\nabla\chi_0\\ \vec{\sigma}\cdot ir_c\nabla\phi_0\end{pmatrix} = \begin{pmatrix}0\\ -2\chi_0\end{pmatrix} \tag{5}$$

where the 0 on the right is a 2×1 null spinor. This zero spinor indicates that the mass energy of the free particle is being ignored, while the effective negative-mass energy of the "vacuum particle" has been doubled. In effect, mass energy for the particle-vacuum system has been conserved by shifting the mass energy of the free particle to the vacuum particle.

The lower of the two simultaneous equations in (5) can be reduced from three to two terms by the assumption

$$\left| ir_c \frac{\partial \chi_0}{c\partial t} \right| \ll |-2\chi_0| \tag{6}$$

if the kinetic energy (from the first equation in (A2)) of the vacuum particle is significantly less than its effective mass energy. Inserting (6) into (5) yields

$$ir_c \frac{\partial}{c\partial t}\begin{pmatrix}\phi_0\\0\end{pmatrix} + \begin{pmatrix}\vec{\sigma}\cdot ir_c\nabla\chi_0\\ \vec{\sigma}\cdot ir_c\nabla\phi_0\end{pmatrix} = \begin{pmatrix}0\\-2\chi_0\end{pmatrix} \tag{7}$$

as the nonrelativistic version of (3). The mass energy of the free particle, and the kinetic energy of the vacuum particle (associated with the lower-left null spinor), are discarded in the Schrödinger approximation.

Separating the two equations in (7) produces

$$ir_c \frac{\partial \phi_0}{c\partial t} + \vec{\sigma}\cdot ir_c\nabla\chi_0 = 0 \tag{8}$$

and

$$\vec{\sigma}\cdot ir_c\nabla\phi_0 = -2\chi_0 \tag{9}$$

and inserting (9) into (8) leads to

$$ir_c \frac{\partial \phi_0}{c\partial t} - \frac{(\vec{\sigma}\cdot ir_c\nabla)^2}{2}\phi_0 = 0\,. \tag{10}$$

Finally, inserting the Pauli-matrix identity (A12)

$$(\vec{\sigma}\cdot ir_c\nabla)^2 = I\,(ir_c\nabla)^2 \tag{11}$$

into (10) yields the free-particle Schrödinger equation

$$ir_c \frac{\partial \phi_0}{c\partial t} = \frac{(ir_c\nabla)^2}{2}\,\phi_0 \quad \text{or} \quad i\hbar\frac{\partial\phi_0}{\partial t} = \frac{(i\hbar\nabla)^2}{2m}\,\phi_0 \tag{12}$$

where the two spin components in ϕ_0 are ignored in this approximation. The scalar harmonic function

$$\phi_0 \longrightarrow \exp\left[-i(Et - \mathbf{p}\cdot\mathbf{r})/\hbar\right] \tag{13}$$

satisfies both equations as it should, and leads to the nonrelativistic energy-momentum relation $E = p^2/2m$, where $\mathbf{p} = m\mathbf{v}$. The equation on the left in (12) expresses the Schrödinger equation in terms of PV gradients.

The vacuum property implied by (11), and the fact that ϕ_0 is a spinor field, show that the vacuum state is a significant (but hidden) part of the nonrelativistic Schrödinger equation. The Dirac-to-Pauli reduction leads to the same conclusion.

3 Minimal coupling

By itself the coupling force (1) is insufficient to split the twofold degeneracy of the spinors in the free-particle Dirac (3) and Schrödinger (12) equations. It takes an external field to effect the split and create the well-known 1/2-spin electron states. The following illustrates this conclusion for the case of the minimal coupling substitution.

The minimal coupling substitution [5, p.78] is

$$p^\mu \longrightarrow p^\mu - eA^\mu/c \tag{14}$$

where e is the magnitude of the observed electron charge, $p^\mu = (E/c, \mathbf{p})$ is the 4-momentum, and $A^\mu = (A_0, \mathbf{A})$ is the electromagnetic 4-potential. Inserting (14) with (A1) and (A2) into the Dirac equation (A3) leads to

$$\left(i\hbar\frac{\partial}{\partial t} - eA_0\right)\psi - c\boldsymbol{\alpha}\cdot\left(\widehat{\mathbf{p}} - \frac{e\mathbf{A}}{c}\right)\psi = mc^2\beta\psi \tag{15}$$

which can be expressed as

$$ir_c \frac{\partial}{c\partial t}\begin{pmatrix}\phi\\\chi\end{pmatrix} + \begin{pmatrix}\vec{\sigma}\cdot(ir_c\nabla + \mathbf{a})\chi\\ \vec{\sigma}\cdot(ir_c\nabla + \mathbf{a})\phi\end{pmatrix} = a_0\begin{pmatrix}\phi\\\chi\end{pmatrix} + \begin{pmatrix}\phi\\-\chi\end{pmatrix} \tag{16}$$

in the 2×1 spinor formulation, where $a_0 \equiv eA_0/mc^2$ and $\mathbf{a} \equiv e\mathbf{A}/mc^2$. Then proceeding as in Section 2 produces

$$ir_c \frac{\partial}{c\partial t}\begin{pmatrix}\phi_0\\0\end{pmatrix} + \begin{pmatrix}\vec{\sigma}\cdot(ir_c\nabla + \mathbf{a})\chi_0\\ \vec{\sigma}\cdot(ir_c\nabla + \mathbf{a})\phi_0\end{pmatrix} = a_0\begin{pmatrix}\phi_0\\\chi_0\end{pmatrix} + \begin{pmatrix}0\\-2\chi_0\end{pmatrix}. \tag{17}$$

The Compton radius in (16) and (17) has been accounted for as a gradient normalizer. The remaining constants (e and m) appear only in association with the 4-potential A^μ — if the external potential vanishes, the electron charge and mass are removed ($a_0 = 0$ and $\mathbf{a} = \mathbf{0}$) from the equations, and (16) and (17) reduce to (3) and (7) respectively. Furthermore, the energy eA_0 appears to increase the energy level of the negative-energy PV continuum. This latter conclusion can be appreciated by combining the two terms on the right side of (17):

$$\begin{pmatrix}\left(eA_0/mc^2\right)\phi_0\\ (eA_0/mc^2 - 2)\chi_0\end{pmatrix} \tag{18}$$

where a_0 has been replaced by its definition. With a constant potential energy $eA_0 = 2mc^2$, the lower parenthesis vanishes and the free-space electron energy and the vacuum-energy spectrum just begin to overlap [1]. This latter result is the phenomenon that leads to the relativistic Klein paradox [5, p. 127].

If it is further assumed that

$$|a_0\chi_0| \ll |-2\chi_0| \tag{19}$$

then (17) becomes

$$ir_c \frac{\partial}{c\partial t}\begin{pmatrix}\phi_0\\0\end{pmatrix} + \begin{pmatrix}\vec{\sigma}\cdot(ir_c\nabla + \mathbf{a})\chi_0\\ \vec{\sigma}\cdot(ir_c\nabla + \mathbf{a})\phi_0\end{pmatrix} = a_0\begin{pmatrix}\phi_0\\0\end{pmatrix} + \begin{pmatrix}0\\-2\chi_0\end{pmatrix} \tag{20}$$

which is the nonrelativistic version of (16). Then eliminating χ_0 from the two simultaneous equations in (20) leads to the equation

$$ir_c\frac{\partial\phi_0}{c\partial t} = \frac{\overrightarrow{\sigma}\cdot(ir_c\nabla+\mathbf{a})\overrightarrow{\sigma}\cdot(ir_c\nabla+\mathbf{a})}{2}\phi_0 + a_0\phi_0 \qquad (21)$$

for the spinor ϕ_0. Equation (21) then leads to the Pauli equation [5, p.81].

Using (A11) to calculate the square of the numerator in the first term on the right of the equal sign in (21) yields

$$ir_c\frac{\partial\phi_0}{c\partial t} = \frac{(ir_c\nabla+\mathbf{a})^2}{2}\phi_0 + \frac{i\overrightarrow{\sigma}\cdot(ir_c\nabla\times\mathbf{a})}{2}\phi_0 + a_0\phi_0 \qquad (22)$$

remembering that ϕ_0 post-multiplies the square before calculation. The first term in (22) contains the electron's orbital angular momentum; and the second its spin, as manifested in the scaler product of $\overrightarrow{\sigma}$ and the curl of the vector potential $\mathbf{A}$. Using (A1), the corresponding spin operator can be expressed as

$$\widehat{s} = \frac{\hbar\overrightarrow{\sigma}}{2} = \frac{e_*^2\overrightarrow{\sigma}}{2c} = \frac{(-e_*)(-e_*)\overrightarrow{\sigma}}{2c} \qquad (23)$$

where one of the charges $(-e_*)$ in (23) belongs to the massive point charge $(-e_*, m)$ and the other to the separate Planck particles $(-e_*, m_*)$ within the PV. The product e_*^2 suggests that the spin may be related to the interaction of the massive point charge with the PV charges when the vacuum is under the influence of a magnetic field $\mathbf{B}\,(=\nabla\times\mathbf{A})$.

4 Conclusions and comments

The physics of the PV state [1, 6] has provided a simple intuitive explanation for the Dirac, Schrödinger, and Pauli equations in terms of the massive point charge $(-e_*, m)$ and its interaction (1) with the PV. It is the ignorance of this coupling force that has obscured the meaning of the Dirac equation since its inception and, as seen in the next paragraph, the meaning of the zitterbewegung frequency.

The electron Compton relation $r_c m = e_*^2$ in (A1) holds for both combinations $(\mp e_*, \pm m)$; so the vacuum hole $(e_*, -m)$ exerts a coupling force on the vacuum state that is the negative of (1). The combination of the two forces explain why the zitterbewegung frequency $(2c/r_c$ [1] [4]) is twice the angular frequency $(mc^2/\hbar = c/r_c)$ associated with the electron mass (from Appendix B).

The purpose of this paper is to illustrate the massive-charge-PV nature of the electron phenomenon; and to reestablish the vacuum state as an essential and necessary part of a complete electron theory, that part that has been superseded by the idea of the quantum field. While the quantum field formalism, like the Green function formalism, is an important tool [5, p. 143] [7], the present author believes that the corresponding quantum field does not constitute an essential physical phenomenon apart from the dynamics of vacuum state (from Appendix C).

Appendix A: Dirac equation

The PV is characterized in part by the two Compton relations [1]

$$r_c mc^2 = r_* m_* c^2 = e_*^2 \quad (= c\hbar) \qquad (A1)$$

connecting the massive point charge $(-e_*, m)$ of the electron to the individual Planck particles $(-e_*, m_*)$ within the degenerate PV, where r_c and m, and r_* and m_* are the Compton radius and mass of the electron and Planck particles respectively. The bare charge $(-e_*)$ is massless and is related to the observed electronic charge $(-e)$ via the fine structure constant $\alpha = e^2/e_*^2$. From (A1), the energy and momentum operators can be expressed as

$$\widehat{E} = i\hbar\frac{\partial}{\partial t} = mc^2\left(ir_c\frac{\partial}{c\partial t}\right) \text{ and } \widehat{\mathbf{p}} = -i\hbar\nabla = mc(-ir_c\nabla) \qquad (A2)$$

the parenthetical factors implying that the operators, operating on the Dirac spinors, provide a measure of the gradients within the PV continuum. In the present free-electron case, these gradients are caused solely by the coupling force (1) and its negative (Appendix B).

The upper and lower limits to the PV negative-energy spectrum are $-mc^2$ and $-m_*c^2$ respectively, where m_* is the Planck mass. The continuum nature of the vacuum is an approximation that applies down to length intervals as small as ten Planck lengths $(10\,r_*)$ or so; that is, as small as $\sim 10^{-32}$cm.

Using (A1) and (A2), the Dirac equation [5, p.74]

$$ie_*^2\frac{\partial\psi}{c\partial t} + \boldsymbol{\alpha}\cdot ie_*^2\nabla\psi = mc^2\beta\psi \qquad (A3)$$

can be expressed as

$$ir_c\frac{\partial\psi}{c\partial t} + \boldsymbol{\alpha}\cdot ir_c\nabla\psi = \beta\psi \qquad (A4)$$

where the 4×4 vector-matrix operator

$$\boldsymbol{\alpha} = \begin{pmatrix} 0 & \overrightarrow{\sigma} \\ \overrightarrow{\sigma} & 0 \end{pmatrix} \qquad (A5)$$

where $\overrightarrow{\sigma} = (\sigma_1, \sigma_2, \sigma_3)$ and

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \qquad (A6)$$

are the three 2×2 Pauli matrices. The 4×4 matrix operator

$$\beta = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} \qquad (A7)$$

where I represents the 2×2 unit matrix and the zeros here and in (A5) are 2×2 null matrices. The covariant gradient operator

$$\frac{\partial}{\partial x^\mu} = \left(\frac{\partial}{\partial x^0}, \frac{\partial}{\partial x^1}, \frac{\partial}{\partial x^2}, \frac{\partial}{\partial x^3}\right) = \left(\frac{\partial}{c\partial t}, \nabla\right) \qquad (A8)$$

is seen in (A2) and (A4) to have its differential coordinates normalized $(\partial x^\mu/r_c)$ by the electron Compton radius.

The 4×1 spinor wavefunction ψ can be expressed as [5, p. 79]

$$\psi = \begin{pmatrix} \phi \\ \chi \end{pmatrix} \tag{A9}$$

where ϕ and χ are the usual 2×1 spinors, and where the two components in each represent two possible spin states. The spinor ϕ is the free-space particle solution and χ is the negative-energy hole solution. Inserting (A9) into (A4), and carrying out the indicated matrix operations, yields the Dirac equation

$$ir_c \frac{\partial}{c\partial t}\begin{pmatrix} \phi \\ \chi \end{pmatrix} + \begin{pmatrix} \overrightarrow{\sigma} \cdot ir_c\nabla\chi \\ \overrightarrow{\sigma} \cdot ir_c\nabla\phi \end{pmatrix} = \begin{pmatrix} \phi \\ -\chi \end{pmatrix} \tag{A10}$$

in terms of the 2×1 spinors.

The following is an important property of the Pauli matrices, and the PV state (because of $\overrightarrow{\sigma}$): the vector Pauli matrix $\overrightarrow{\sigma}$ obeys the identity [5, p.12]

$$(\overrightarrow{\sigma} \cdot \mathbf{a})(\overrightarrow{\sigma} \cdot \mathbf{b}) = I\,\mathbf{a} \cdot \mathbf{b} + i\overrightarrow{\sigma} \cdot \mathbf{a} \times \mathbf{b} \tag{A11}$$

where $\mathbf{a}$ and $\mathbf{b}$ both commute with $\overrightarrow{\sigma}$, but are otherwise arbitrary three-vectors. Using (A11) (with $\mathbf{a} = \mathbf{b} = r_c\nabla$) leads to

$$(\overrightarrow{\sigma} \cdot r_c\nabla)^2 = I\,(r_c\nabla)^2 \tag{A12}$$

which connects the normalized ∇ operator in the relativistic Dirac equation to the same operator in the nonrelativistic Schrödinger equation.

Inserting the operators from (A2) into (A10) and rearranging the result leads to the two simultaneous equations

$$\left(\widehat{E} - mc^2\right)\phi = c\overrightarrow{\sigma} \cdot \widehat{p}\chi \tag{A13}$$

and

$$\left(\widehat{E} + mc^2\right)\chi = c\overrightarrow{\sigma} \cdot \widehat{p}\phi\,. \tag{A14}$$

Then, pre-multiplying (A13) by $(\widehat{E} + mc^2)$ and using (A14) and (A11) leads to

$$\left(\widehat{E^2} - m^2c^4\right)\phi = c^2\widehat{p^2}\,\phi \tag{A15}$$

and, after reversing the process, to an identical equation for χ. Thus both ϕ and χ separately obey the Klein-Gordon equation [5, p.31].

Appendix B: Zitterbewegung frequency

The following rough heuristic argument identifies the two coupling forces that explain why the zitterbewegung frequency [1, 4] is twice the angular frequency ($mc^2/\hbar = c/r_c$) associated with the electron mass energy.

The force the massive point charge $(-e_*, m)$ exerts on the PV is given by equation (1) which, using $r = r_c + \Delta r$ and $r_c = e_*^2/mc^2$, leads to

$$\frac{e_*^2}{(r_c + \Delta r)^2} - \frac{mc^2}{r_c + \Delta r} = -\frac{(e_*^2/r_c^3)\Delta r}{(1 + \Delta r/r_c)^2} \approx -\left(\frac{e_*^2}{r_c^3}\right)\Delta r \tag{B1}$$

for small $\Delta r/r_c$. This yields the harmonic oscillator motion from Newton's second law

$$\frac{d^2\Delta r}{dt^2} = -\left(\frac{e_*^2}{mr_c^3}\right)\Delta r = -\left(\frac{c}{r_c}\right)^2\Delta r \tag{B2}$$

with the "spring constant" (e_*^2/r_c^3) and oscillator frequency c/r_c. The corresponding motion that is due to the vacuum hole $(e_*, -m)$ (whose charge and mass fields exert a force that is the negative of (1)) is

$$-\frac{d^2\Delta r}{dt^2} = +\left(\frac{c}{r_c}\right)^2\Delta r \tag{B3}$$

showing that the massive free charge and the vacuum hole cause identical accelerations within the PV continuum.

The total vacuum acceleration is the sum of (B2) and (B3)

$$\frac{d^2\Delta r}{dt^2} = -2\left(\frac{e_*^2}{mr_c^3}\right)\Delta r = -2\left(\frac{c}{r_c}\right)^2\Delta r \tag{B4}$$

with the corresponding harmonic oscillator frequency

$$\sqrt{\frac{2e_*^2}{mr_c^3}} = \sqrt{2}\,\frac{c}{r_c} \tag{B5}$$

which is $\sqrt{2}$ times the angular frequency associated with the electron mass energy. Given the roughness of the calculations, this result implies that the combined massive-charge forces, acting simultaneously on the PV continuum, are the source of the zitterbewegung with its $2c/r_c$ frequency.

Appendix C: Quantum field

The PV is envisioned as a *degenerate* negative-energy sea of fermionic Planck particles. Because of this degeneracy, the vacuum experiences only small displacements from equilibrium when stressed. Thus the displacements due to the coupling force (1) are small, and so the potential energy corresponding to the stress can be approximated as a quadratic in those displacements. This important result enables the vacuum to support normal mode coordinates and their assumed quantum fields, as explained in the simple demonstration to follow.

The normal mode connection [8, pp. 109–119] to the quantum field can be easily understood by examining a string, stretched between two fixed points in a stationary reference frame, that exhibits small transverse displacements from equilibrium. In this case, the corresponding potential energy can be expressed in terms of quadratic displacements. If the displacements are represented by the function $\phi(t, x)$ at time t and position x along the string, then the quadratic assumption implies that the displacements must obey the wave equation

$$\frac{1}{c^2}\frac{\partial^2\phi}{\partial t^2} = \frac{\partial^2\phi}{\partial x^2} \tag{C1}$$

where c is a propagation velocity. The string geometry leads to the Fourier series representation

$$\phi(t, x) = \sum_{n=1}^{N} a_n(t) \sin(n\pi x/L) \tag{C2}$$

for the standing wave on the string, where L is the string length. Inserting (C2) into (C1) produces

$$\ddot{a}_n(t) = -\omega_n^2 a_n(t) \qquad \text{where} \qquad \omega_n = n\pi c/L \tag{C3}$$

and where the amplitude $a_n(t)$ is that of a harmonic oscillator.

The constant characterizing the Dirac equation is the Compton radius r_c. So it is reasonable to set the string length $L \sim r_c$ to determine the fundamental frequency $\omega_1 = \pi c/L$ in (C3). Furthermore, the harmonics of ω_1 can have wavelengths of the order of the Planck length r_* (antiparticle excitation is, of course, ignored in this rough argument); so the length L can be subdivided

$$N = \frac{L}{\text{minimum length division of string}} \sim \frac{r_c}{r_*}$$

$$= \frac{3.86 \times 10^{-11}}{1.62 \times 10^{-33}} \sim 10^{22} \tag{C4}$$

times, and ϕ in (C2) can be expressed as an integral if convenient since $r_* \ll r_c$.

The total energy of the vibrating string can thus be expressed as

$$E = \int_0^L \left[\frac{\rho}{2} \left(\frac{\partial \phi}{\partial t} \right)^2 + \frac{\rho}{2} c^2 \left(\frac{\partial \phi}{\partial x} \right)^2 \right] dx \tag{C5}$$

which, inserting (C2) into (C5), results in [8, p.117]

$$E = \frac{L}{2} \sum_{n=1}^{N} \left[\frac{\rho \dot{a}_n^2}{2} + \frac{\rho \omega_n^2 a_n^2}{2} \right] \tag{C6}$$

where the first and second terms in (C5) and (C6) are the kinetic and potential string energies respectively (ρ is the string density).

The crucial significance of (C6) is that it is a sum of independent normal-mode energies, where the $a_n(t)$ are the normal mode coordinates. From this normal mode setting, the quantum field energy

$$E = \sum_{n=1}^{N} \left(n_n + \frac{1}{2} \right) \hbar \omega_n = mc^2 \sum_{n=1}^{N} \left(n_n + \frac{1}{2} \right) r_c k_n \tag{C7}$$

is defined, where n_n is the number of normal modes associated with the wavenumber $k_n = \omega_n/c$. In effect, the integers $n_n (\geqslant 0)$ determine the quantized energy level of each normal mode oscillator $a_n(t)$. The $1/2$ component in (C7) is the zero-point energy of the string-vacuum system.

At this point the quantum-field formalism discards the preceding foundation upon which the fields are derived, and assumes that the fields themselves are the primary reality [8, p. 119]. Part of the reason for this assumption is that, in the past, no obvious foundation was available. However, the demonstration here provides such a foundation on the simple, but far-reaching assumption that the vacuum is a degenerate state which can sustain a large stress without a correspondingly large strain.

Submitted on September 11, 2013 / Accepted on September 21, 2013

References

1. Daywitt W.C. The Electron-Vacuum Coupling Force in the Dirac Electron Theory and its Relation to the Zitterbewegung. *Progress in Physics*, 2013, v. 3, 25.
2. Puthoff H.E. Gravity as a Zero-Point-Fluctuation Force. *Physical Review A*, 1989, v. 39, no. 5, 2333–2342.
3. Daywitt W.C. The Source of the Quantum Vacuum. *Progress in Physics*, 2009, v. 1, 27.
4. Barut A.O. and Bracken A.J. *Zitterbewegung* and the Internal Geometry of the Electron. *Physical Review D*, 1981, v. 23, no. 10, 2454–2463.
5. Gingrich D.M. Practical Quantum Electrodynamics. CRC, The Taylor & Francis Group, Boca Raton, London, New York, 2006.
6. Daywitt W.C. The Planck Vacuum. *Progress in Physics*, 2009, v. 1, 20. See also www.planckvacuum.com.
7. Milonni P.W. The Quantum Vacuum—an Introduction to Quantum Electrodynamics. Academic Press, New York, 1994.
8. Aitchison I.J.R., Hey A.J.G. Gauge Theories in Particle Physics Vol. 1. Taylor & Francis, New York, London, 2003.

Geometric Distribution of Path and Fine Structure

Janez Špringer

Cankarjeva Cesta 2, 9250 Gornja Radgona, Slovenia, EU. E-mail: info@lekarna-springer.si

Previously (*Progr. Phys.*, v.2, 105–106) one predicted the exact value of the inverse fine structure constant respecting the double surface concept on Bohr orbit. In this paper one extends the same principle on the geometric distribution of the frequencies of the path of the electron in the ground state of Hydrogen atom. The inverse fine structure constant reflects the kind of the distribution and the later increases the constant in the range of the fifth decimal from $\alpha^{-1}_{0-sided} > 137.036006$ to $\alpha^{-1}_{\infty-sided} < 137.036018$.

1 Theoretical background

The number 137 expresses the translation component n of the path s of the electron on Bohr orbit [1]. Let us consider other translations n around this value are also possible. Each of them has its own frequency:

$$f_z = f(z), \text{ where } z = n - 137,\ n \in \mathbb{Z}. \tag{1}$$

It is also reasonable to assume the sum of the frequencies of all translations equals the unit which is the frequency of the whole translation of the path:

$$F_z = \sum f_z = 1. \tag{2}$$

The two-sided distribution ranges from the translation $n = -\infty$ to $n = 137$ on Bohr orbit and further from there to $n = \infty$. Overall interval is opened since the frequencies at the infinite ends $f_{\pm\infty}$ equal zero and can be ignored. There are also possible even-sided distributions provided on the arbitrary number of two-sided dimensions. From this point of view the non-distribution at $n = 137$ is regarded as zero-sided.

Each translation n belongs to its path s_n so the frequency of the former is identical to the frequency of the later. Product of the given frequency of the path f_z and the path s_n itself is the pondered partial path $f_z \cdot s_n$ inside the whole distribution of the path:

$$s_{whole} \cdot F_z = s_{whole} = \sum f_z \cdot s_n\,. \tag{3}$$

The inverse fine structure constant reflecting the whole distribution of the path [2] can be then expressed as:

$$\alpha^{-1}_{distributed} = \sum f_z \cdot s_n. \tag{4}$$

According to the double-surface concept [2] the value of the path s_n depends on the translation n:

$$s_n = n\left(2 - 1\Big/\sqrt{1 + \pi^2/n^2}\right), \text{ where } n \in \mathbb{Z}. \tag{5}$$

Knowing the type of the distribution function of frequencies $f(z)$ the inverse fine structure constant $\alpha^{-1}_{distributed}$ can be calculated. And vice versa, knowing the inverse fine structure constant $\alpha^{-1}_{distributed}$ the type of the distribution function of frequencies $f(z)$ can be speculated.

Our subject of interest in this paper is the geometric distribution of the frequencies of the path with ratio $1/2$ where the jumping of the electron to the non-adjacent positions is not allowed.

2 The two-sided geometric distribution

This is the symmetric distribution of the frequencies of the path provided on and around the zero numbered position z at $n = 137$:

$$f_z = \frac{1}{3}\frac{1}{2^{|z|}}, \text{ where } z \in \mathbb{Z}. \tag{6}$$

The sum of the frequencies f_z of all translations n from $-\infty$ to $+\infty$ equals the unit:

$$F_z = \sum_{z=-\infty}^{z=\infty} f_z = 1,$$

since

$$\sum_{z=-\infty}^{z=\infty} \frac{1}{3}\frac{1}{2^{|z|}} = \frac{1}{3}\sum_{z=-\infty}^{z=-1}\frac{1}{2^{|z|}} + \frac{1}{3}\sum_{z=0}^{z=0}\frac{1}{2^z} + \frac{1}{3}\sum_{z=1}^{z=\infty}\frac{1}{2^z} = \\ = \frac{1}{3} + \frac{1}{3} + \frac{1}{3} = 1. \tag{7}$$

The value of the inverse fine structure constant reflecting the 2-sided geometric distribution of the frequencies of the path of the electron in the ground state of Hydrogen atom can be calculated with the help of equations (1), (4), (5) and (6):

$$\alpha^{-1}_{2-sided} = \frac{1}{3}\sum_{n=-\infty}^{n=\infty} \frac{n\left(2 - 1\Big/\sqrt{1 + \pi^2/n^2}\right)}{2^{|n-137|}}\,. \tag{8}$$

The values of the frequencies of the path f_z rapidly lessen in the negative as well as positive direction from the zero numbered position z on Bohr orbit so the enough accurate value of the constant can be calculated numerically on the appropriate finite interval, for instance $n = [104, 170]$:

$$\alpha^{-1}_{2-sided} \approx 137.036014. \tag{9}$$

3 The even-sided geometric distribution

On the arbitrary number of sides generalized distribution of the frequencies of the path makes sense to be taken into account when some extra two-sided dimensions are proposed to be involved. The distribution on the even number of k sides is then expressed as:

$$f_z = \frac{1}{k+1}\frac{1}{2^{|z|}}, \text{ where } z \in \mathbb{Z} \text{ and } k = 2m,\ m \in \mathbb{N}_0. \quad (10)$$

The sum of the frequencies f_z of the infinite number of translations n on all k-sides and one zero position equals the unit:

$$F_z = \sum_{z=-\infty}^{z=\infty} f_z = 1,$$

since

$$\sum_{z=-\infty}^{z=\infty}\frac{1}{k+1}\frac{1}{2^{|z|}} = \frac{k}{2}\sum_{z=-\infty}^{z=-1}\frac{1}{k+1}\frac{1}{2^{|z|}} + \sum_{z=0}^{z=0}\frac{1}{k+1}\frac{1}{2^{z}} + \\ + \frac{k}{2}\sum_{z=1}^{z=\infty}\frac{1}{k+1}\frac{1}{2^{z}} = \frac{k}{2}\frac{1}{k+1} + \frac{1}{k+1} + \frac{k}{2}\frac{1}{k+1} = 1. \quad (11)$$

The value of the inverse fine structure constant reflecting the k-sided geometric distribution of the path of the electron in the ground state of Hydrogen atom is found with the help of equations (1), (4), (5), (10) and (11):

$$\alpha^{-1}_{k-sided} = \frac{k}{2(k+1)}\sum_{n=-\infty}^{n=136}\frac{n\left(2-1\Big/\sqrt{1+\pi^2/n^2}\right)}{2^{|n-137|}} + \\ + \frac{137\left(2-1\Big/\sqrt{1+\pi^2/137^2}\right)}{k+1} + \\ + \frac{k}{2(k+1)}\sum_{n=138}^{n=\infty}\frac{n\left(2-1\Big/\sqrt{1+\pi^2/n^2}\right)}{2^{n-137}}, \quad (12)$$

where $n \in \mathbb{Z}$ and $k = 2m$, $m \in \mathbb{N}_0$.

The enough accurate value of the constant can be calculated numerically on the appropriate finite interval. For the acceptable results rounded on the six decimals can be used the finite intervals $n = 137 \pm 33$ instead of the infinite ones $n = 137 \pm \infty$. There is the infinite number of the even-sided distributions available from $k = 0$ to $k = \infty$. The 2-sided distribution at $k = 2$ is only one of them.

4 The non-distribution

Such special distribution of the frequencies of the path of the electron is considered on the zero position and zero sides on Bohr orbit. At $k = 0$ the equation (10) and (11) are simplified to $f_z = F_z = 1$ so the equation (12) takes the known form useful for the calculation of the theoretical inverse fine structure constant[2],(5):

$$\alpha^{-1}_{0-sided} = 137\left(2-1\Big/\sqrt{1+\pi^2/137^2}\right) > 137.036006. \quad (13)$$

5 The infinite-sided geometric distribution

Such special distribution of the frequencies of the path of the electron takes place on the infinite sides around Bohr orbit. At $k = \infty$ the equation (12) is shortened for the vanished middle term and transformed into the next simplified form useful for the finding the theoretical inverse fine structure constant:

$$\alpha^{-1}_{\infty-sided} = \sum_{n=-\infty}^{n=136}\frac{n\left(2-1\Big/\sqrt{1+\pi^2/n^2}\right)}{2^{|n-137|+1}} + \\ + \sum_{138}^{n=\infty}\frac{n\left(2-1\Big/\sqrt{1+\pi^2/n^2}\right)}{2^{n-137+1}} < 137.036018. \quad (14)$$

6 The inverse fine structure reflecting the geometric distribution

The distributed value of the inverse fine structure constant seems to be greater than the non-distributed one since:

$$\alpha^{-1}_{\infty-sided} \approx 137.036018 > \alpha^{-1}_{2-sided} \approx \\ \approx 137.036014 > \alpha^{-1}_{0-sided} \approx 137.036006. \quad (15)$$

The answer doesn't lie in the frequency of the path f_z which otherwise equally decreases on both sides of the number 137 but depends on the factor $\left(2-1\Big/\sqrt{1+\pi^2/n^2}\right)$ which increases more with $n < 137$ than decreases with $n > 137$. The overall effect is thus the increasing value of the distributed inverse fine structure constant inside the range of the fifth decimal.

7 Conclusions

According to the double surface concept the exact inverse fine structure constant reflects the kind of the distribution of the frequencies of the path of the electron in the ground state of Hydrogen atom. The factor $\left(2-1\Big/\sqrt{1+\pi^2/n^2}\right)$ asymmetrically changes partial values of the constant what results the increasing value of the whole constant. The number of sides of the distribution influences the above change in the range of the fifth decimal. The zero-, two- and infinite-sided geometric distribution of the frequencies of the path yields on the six decimal rounded inverse fine structure constant of 137.036006, 137.036014 and 137.036018, respectively.

Dedication

This fragment is dedicated to my daughters Alenka, Manica and Natalija.

Submitted on September 17, 2013 / Accepted on September 22, 2013

References

1. Špringer J. Fine structure constant as a mirror of sphere geometry. *Progress in Physics*, 2013, v. 1, 12–14.
2. Špringer J. Double surface and fine structure. *Progress in Physics*, 2013, v. 2, 105–106.

A Prediction of an Additional Planet of the Extrasolar Planetary System Kepler-62 Based on the Planetary Distances' Long-Range Order

Felix Scholkmann

Bellariarain 10, 8038 Zürich, Switzerland. E-mail: felix.scholkmann@gmail.com

Recently, the discovery of the extrasolar planetary system Kepler-62 comprising five planets was reported. The present paper explores whether (i) the sequence of semi-major axis values of the planets shows a long-range order, and whether (ii) it is possible to predict any additional planets of this system. The analysis showed that the semi-major axis values of the planets are indeed characterized by a long-range order, i.e. the logarithmic positions of the planets are correlated. Based on this characteristic, an additional planet at 0.22 AU in the Kepler-62 system is predicted.

1 Introduction

In April 2013, NASA's Kepler Mission reported [1] the detection of an extrasolar planetary system comprising five planets (Kepler-62b, c, d, e and f) orbiting a star (Kepler-62) of spectral type K2, luminosity class V, 69 ± 0.02% the mass and 63 ± 0.02% the radius of the Sun. The Kepler-62 extrasolar planetary system is located in the constellation Lyra, ∼1200 light years away from Earth. The five planets have a size of 1.31, 0.54, 1.95, 1.61 and 1.41 Earth radii ($R_\oplus$). The two outermost planets (e, f) are likely to be solid planets possibly with liquid water on their surfaces since their position is within Kepler-62's Habitable Zone. The five planets were detected by analyzing the brightness variations of Kepler-62 based on images obtained by the Kepler spacecraft.

In an analysis of distances between planets of our solar system (including the dwarf planet Pluto and the asteroid Ceres) it was shown by Bohr and Olsen [2] that the sequence of distances show a long-range order on a logarithmic scale, i.e. the logarithmic positions of the planets are correlated and follow a periodic pattern; they seem to obey a "quantization". The authors tested the statistical significance of the obtained long-range order by using a permutation test, which revealed that the regularity of the distances between the planets in our solar system is very unlikely to have originated by chance.

In a subsequent study by the same authors [3], they applied their analysis to the extrasolar planetary system HD 10180 and determined that (i) the logarithmic position of the six planets show also a long-range order, and (ii) that this property is enhanced when including a seventh (hypothetically existing) planet at a position of 0.92 ± 0.05 AU, i.e. between the planets HD 10180f and HD 10180g. Based on this analysis, they postulated a possible additional planet in the HD 10180 system at a distance of 0.92 AU.

The goal of the present analysis was to apply the same data analysis approach [2, 3] to the recently discovered extrasolar planetary system Kepler-62 and thus to analyze whether (i) the semi-major axis values of the planets show a long-range order, and whether (ii) the analysis predicts additional planets of this system.

2 Materials and methods

2.1 Data

The parameter values of the Kepler-62's exoplanets were obtained from the listing in Borucki et al. [1]. In particular, two parameters were selected for the present analysis: the semi-major axis (a) and the radius (r) of each planet. For the values, see Table 1.

Planet	i	a [AU]	a [km]	r [$R_\oplus$]	r [km]	$\hat{a}$
62b	1	0.0553	8.2728×10^6	1.31	8355	2.1130
62c	2	0.0929	1.3898×10^7	0.54	3444	2.6317
62d	3	0.120	1.7952×10^7	1.95	12437	2.8877
62e	4	0.427	6.3878×10^7	1.61	10269	4.1570
62f	5	0.718	1.0741×10^8	1.41	8993	4.6767

Table 1: Kepler-62 system parameters according to [1]. i: planet number counting outwardly from the star Kepler-62, a: semi-major axis, r: radius of the planet, ($\hat{a}_i = \ln(a_i/10^6\,\text{km})$), a and r are given in two different units ([AU], [km]) and ([$R_\oplus$], [km]), respectively.

2.2 Data analysis

For the analysis, the semi-major axis value (given in units of 10^6 km) of each exoplanet was first divided by 10^6 km, then logarithmized ($\hat{a}_i = \ln(a_i/10^6\,\text{km})$) and according to these values a multimodal probability distribution function (PDF) $p(\hat{a})$, as introduced by Bohr and Olsen [2], was calculated by

$$p(\hat{a}) = \sum_{i=1}^{N} \alpha_i \, e^{-\beta}, \tag{1}$$

with $N = 5$ (i.e. the maximum number of planets of Kepler-62) and β given as

$$\beta = \frac{j - \hat{a}_i}{w_p / 2\sqrt{2\ln(2)}}, \tag{2}$$

for $j = 1, 1.01, 1.02, \ldots, 10$, with w_p the width (i.e. the full-width-at-half-maximum) of each Gaussian peak of the PDF, and α_i a scale factor. The scale factor defined the magnitude

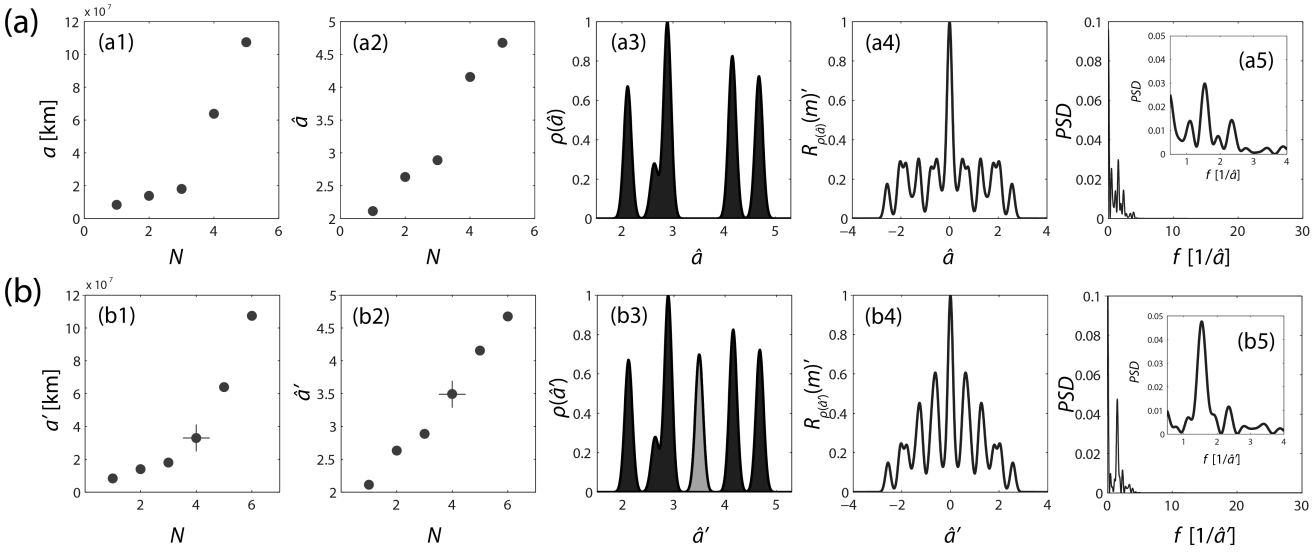

Fig. 1: Results of the analysis of the multimodal PDF $\rho(\hat{a})$ (a1-a5) and the new one $\rho(\hat{a}')$ with the additional hypothetical exoplanet (marked with a cross in Fig. (b1) and (b2), and marked with a black are of the Gaussian peak in Fig. (b3)) found using the optimization approach visualized in Fig. 2.

of each peak. For the present analysis, the scale factor was assigned to the radius of the specific planet, i.e. $\alpha_i = r_i$. The rationale for this definition is that larger planets should be contribute more to the overall multimodal PDF than smaller planets. A linear relationship was chosen rather than the non-linear one used by Bohr and Olsen [2, 3] in order to circumvent the definition of the specific type of non-linear relationship which is unknown per se. For the width of each peak, $w_p = 0.25$ was used which ensures an optimum compromise between a too strong overlap of the Gaussian peaks on the one side and to small peaks on the other. Thus, $\rho(\hat{a})$ represents a sum of Gaussian peaks located at the logarithmized planets semi-major axis values ($\hat{a}$) and weighted by (α_i), the individual radius value of the planet.

In the next step, the autocorrelation sequence of the multimodal PDF was calculated according to

$$R_{\rho(\hat{a})}(m) = \sum_{n=0}^{N-m-1} \rho(\hat{a}_{n+m})\,\rho(\hat{a}_n), \qquad (3)$$

for $m = 1, 2, \ldots, 2N-1$, with N the number of samples of $\rho(\hat{a})$. Then, the autocorrelation function (ACF) was determined by

$$R_{\rho(\hat{a})}(m)' = \frac{1}{R_{\rho(\hat{a})}(1)}\, R_{\rho(\hat{a})}(m), \qquad (4)$$

i.e. $R_{\rho(\hat{a})}(m)$ was normalized by its maximum value given by $R_{\rho(\hat{a})}(1)$ so that $R_{\rho(\hat{a})}(1)' = 1$. The type and grade of the order (short- or long-range) of the input sequence can be determined using the ACF characteristics.

In order to quantify the periodicity in the ACF (i.e. the long-range order of the input sequence), in the next step the frequency-dependent power spectral density (PSD), i.e. the power spectrum (PS), of the multimodal PDF $\rho(\hat{a})$ was calculated by the periodogram method, which is the windowed discrete Fourier transform (DFT) of the biased estimate of autocorrelation sequence. For the calculation, 2^{12} points in the DFT were used by zero-padding $\rho(\hat{a})$ to a length of 2^{12} enabling a proper frequency resolution.

In order to analyze whether an additional hypothetical planet increases the long-range order, the above-mentioned signal processing steps (i.e. calculation of the multimodal PDF, the ACF and the PS) were repeated with the input signal $\rho(\hat{a})$ in which an additional Gaussian peak was inserted, corresponding to the hypothetical exoplanet's position. The high of the peak was set to the mean values of the radius of the five exoplanets. The new peak was introduced between the peaks associated with values of Kepler-62e and Kepler-62f since visual inspection reveals a gap in the multimodal PDF in this region. The semi-major axis value was varied between 0.15-0.38 AU and the corresponding ACF and PS were calculated. For each PS, the maximum PSD value of the fundamental frequency of $\rho(\hat{a})$ (i.e. the first peak after the global maximum at position 0) was calculated. From the obtained values, the maximum was determined which indicate the strongest long-range order of the corresponding sequence with the added new exoplanet. This new multimodal PDF was denoted as $\rho(\hat{a}')$, with $\hat{a}'$ the vector with the new semi-major axis values.

3 Results

The analysis of the semi-major axis values of Kepler-62's planets b-f revealed an exponential like function (Fig. 1(a1))

or a quasi linear one when logarithmized values were used (Fig. 1(a2)).

The calculated multimodal PDF is shown in Fig. 1(a3). The ACF and the PS are shown in Fig. 1(a3) and 1(a4), respectively. The search of the optimal semi-axis value of the additional (hypothetical) planet revealed that a global maximum of the PSD value in the frequency range of 1.1538 $1/\hat{a}$ (≈ 0.6502 units of $\hat{a}$) can be clearly determined, as depicted in Fig. 2. Thus, the analysis predicts an additional planet at a distance of 0.22 AU from the star Kelper-62. The characteristics of the resulting new multimodal PDF $\rho(\hat{a}')$ with all six planets are shown in Fig. 1(b1-b5).

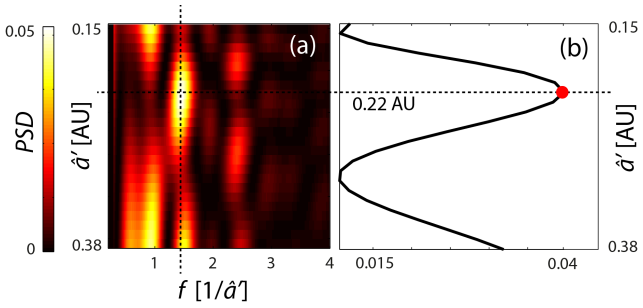

Fig. 2: (a) Color-coded visualization of the PSD values for the $\rho(\hat{a})$ function with the added hypothetical exoplanet at different positions. (b) Function of the PSD values for the frequency of 1.1538 $1/\hat{a}$. The global maximum indicates the value which corresponds to the strongest increase in the long-range order.

4 Discussion and conclusion

From the analysis conducted in the present study, the following conclusions can be drawn:

(i) The positions of the exoplanets Kepler-62a–f show a long-range order inferred from the peak-like structure (four peaks) in the ACF which is captured by the power spectrum as one single peak, corresponding to linear periodicity of the logarithmized distances between the planets.

(ii) The strength of the long-range order increases when an additional planet with a distance of 0.22 AU from the star is added to the five observed ones. This result was obtained by an optimization procedure testing all possible positions for this planet in the gap between Kepler-62e and Kepler-62f.

A prediction of possible additional planets in the Kepler-62 extrasolar system was put forward also recently by Bovaird and Lineweaver [4]. They applied a two-parameter fit to 68 different extrasolar planetary systems in total and predicted 141 additional planets. For the fitting they used a function (denoted by them as a modified Titius-Bode relation) of the form $a_n = \alpha C^n$, with a_n an the semi-major axis, two free parameters (α, C), and n a variable with the quantized values $n = 1, 2, 3, \ldots$. Based on their approach, they predicted for the Kelpler-62 system 7 additional planets with semi-major distance values of 0.07, 0.15, 0.20, 0.26, 0.33, 0.55 and 0.92 AU. Thus, the approach of Bovaird and Lineweaver predicts a finer periodicity compared to the prediction (0.22 AU) described by the present paper. Only the future will tell which approach is better in modeling the exoplanetary characteristics, i.e. the next discovery of an exoplanet of Kepler-62.

By the best of my knowledge, the two predictions (by Bovaird and Lineweave, and the present one), are the only ones published at the present concerning the extrasolar planetary system Kepler-62.

For other extrasolar planetary systems, various authors have reported a periodicity/quantization of the planetary positions and predicted additional orbits/planets based on this.

For example, Naficy et al. [5], recently compared two approaches for modelling and predicting by using either a squared model of the form $r_n = GM\, n^2/(v_0^2\, k^2)$ (with r_n the orbital radius of the n-th planet, G the gravitational constant, M the mass of a central body of the system, and the free parameters v_0^2, k, and n) or an exponential one given by $r_n = a\, e^{bn}$ (with a, b, n free parameters). In both cases, the parameter values of n are integers. The authors concluded that the "exponential model has a better coincidence to observational data" [5]. In addition they observed a relation between the values of the b parameter and the mass of the central star of the system, indicating a possible physical mechanism underlying the exponential model. The squared model was also used in a study analyzing extrasolar planetary systems conducted by Rubčić and Rubčić [11].

Another study based on an exponential model was conducted by Poveda and Lara [24] to examine the extrasolar planetary system 55 Cancri. However, problems with this study were pointed out later [23].

In another study, Panov [6] applied an exponential model of the type $a_n = C\, e^{2n/k}$ to extrasolar planetary systems and reported a good fit as well as predictions of additional planets.

As early as 1996, Nottale found that "the distribution of the semi-major axis of the firstly discovered exoplanets was clustered around quantized values according to the law $a/GM = (n/w_0)^2$, in the same manner and in terms of the same constant $w_0 = 144$ km/s as in our own inner Solar System" [7, 8]. This approach is a result of the "scale relativity" theory developed by Nottale [9, 10, 32, 33]. In 2008, an updated analysis involving 300 exoplanets was published [10] which confirmed and extended the validity of the initial analysis of 1996.*

An analysis with 443 exoplanets (i.e. all known in 2011) was conducted by Zoghbi [26]. This revealed a quantization of the planet's angular momentum which was shown to have a probability of $p < 0.024$ being due to pure chance.

*It would be worthwhile and interesting to repeat the analysis with the presently 732 confirmed exoplanets (September 2013, http://exoplanets.org).

In another study, using the equation $r_n = GM/(c\,\alpha\,e^n)$, with α the dimensionless fine structure constant of $\sim 1/137$ and c the speed of light, Pintr et al. [12] reported a strong agreement between the orbital data of the two analyzed extrasolar planetary system and the expected values. The interesting thing about this work is that the equation is derived from a physical theory describing the effects of electric and magnetic effects on the evolution of a solar system.

Finally, as mentioned earlier, employing a similar method to the one used in this paper (i.e. analysis of correlation property of the logarithmized planetary positions), Olsen and Bohr [3] analysed the extrasolar planetary HD 10180 and predicted an additional planet at 0.92 ± 0.05 AU.

Apart from analyzing extrasolar planetary systems, empirical relationships for the distances of the planets of our solar system started to be published centuries ago when J. D. Titius (1729–1796) and E. Bode (1744–1826) described an apparent regularity of the planetary radii, later known as the Titius-Bode law (expressed in 1787 in its more modern mathematical form by Wurm: $r_n = 0.4 + 0.3 \times 2^n$, $n = -\infty$ (Mercury), $0, 1, 2, \ldots$) [13]. This equation predicted the position of Uranus, but failed to fit for the planetary positions of Neptune and Pluto. Based on the many studies about regularities in planetary distances/radii conducted until now, the Titius-Bode law can be regarded as a first phenomenological description of a possible fundamental law of planetary spacing. The work of Bohr and Olsen [2, 3] in particular strongly suggests that the orbital spacing of planetary systems obey a long-range order and not a simple short-range one, supporting the notion that the quantization is not down to chance.

Concerning the physical mechanism involved in creating a long-range order in planetary systems, this issue is not resolved yet. However, important approaches have been put forward over the last decades. For example, Wells showed that the planetary distances can be "accurately predicted by the eigenvalues of the Euler-Lagrange equations resulting from the variation of the free energy of the generic plasma that formed the Sun and planets" [14, 15]. Further research of the author led him to conclude that "a unification of the morphology of the solar system" and other astrophysical phenomena "can be accomplished by a basic consideration of the minimum-action states of cosmic and/or virtual vacuum field plasmas" [16]. Finally he came to the conclusion that a unification of all physical forces can be derived based on the assumption that they are regarded "as 'fluid' or 'Magnus' forces generated by vortex structures (particles) in the virtual plasma gas" [15–17]. The work of Wells should be carefully reconsidered since it might be a key to understanding regular patterns, long-range orders and quantization of astronomical systems and structures. In addition, analysis based on the theoretical framework of stochastic electrodynamics (SED) that shed new light on the origin of the solar system [18], and also the finding of Graner and Dubrulle [19, 20] that Titius-Bode-like laws appear when assuming a scale and rotational invariance of the protoplanetary system, might also be important for an understanding of the observed patterns.

Other approaches worth exploring for further research are that based on large-scale quantization in space plasmas [22], modelling celestial mechanics using the Schrödinger equation [21, 27, 29, 39–41, 43], resonance effects [25, 28], orbital angular momentum quantization per unit mass [30, 37], fractal scaling modeling using the continued fraction method [31], conservation of mass and momentum, and stability of the angular momentum deficit [35, 36], the stochastisation hypothesis [34], macroscopic quantization due to finite gravitational propagation speed [38], and the Weyl-Dirac approach to gravity [42].

One significant difficulty in explaining the observed regularities of distances is the fact that planets can migrate large distances after their formation (e.g. [44–48]). A model that gives an explanation of the regularities must include this observed fact. One possible explanation might be to regard the quantization pattern as an attractor in a phase-space of the planet's migration movements.

In conclusion, the present analysis of the extrasolar planetary system Kepler-62 reveals that (i) the semi-major axis values of the planets show a long-range order, and (ii) that there might be an additional planet at the distance of 0.22 AU between Kepler-62e and Kepler-62f.

Submitted on September 18, 2013 / Accepted on September 29, 2013

References

1. Borucki W. J., Agol E., Fressin F., Kaltenegger L., Rowe J., et al. Kepler-62: A five-planet system with planets of 1.4 and 1.6 Earth radii in the Habitable Zone. *Science*, 2013, v. 340 (6132), 587–590.
2. Bohr J., and Olsen K. Long-range order between the planets in the Solar system. *Monthly Notices of the Royal Astronomical Society*, 2010, v. 403, L59–L63.
3. Olsen K. and Bohr J. Pair-correlation analysis of HD 10180 reveals a possible planetary orbit at about 0.92 AU. 2010, arXiv: astro-ph.EP/1009.5507.
4. Bovaird T. and Lineweaver, C.H. Exoplanet predictions based on the generalised Titius-Bode relation. 2013, arXiv: astro-ph.EP/1304.3341.
5. Naficy K., Ayubinia A. and Saeedi M. Exponential law as a more comp atible model to describe orbits of planetary systems. *Iranian Journal of Physics Research*, 2012, v. 12 (3), 25–31.
6. Panov K. P. The orbital distances law in planetary systems. *The Open Astronomy Journal*, 2009, v. 2, 92–94.
7. Nottale L., Schumacher G. and Lefévre E.T. Scale-relativity and quantization of exoplanet orbital semi-major axes. *Astronomy and Astrophysics*, 2000, v. 361, 379–387.
8. Nottale L. Scale-relativity and quantization of extra-solar planetary systems. *Astronomy and Astrophysics*, 1996, v. 315, L9–L12.
9. Nottale L. Fractal Space-Time and Microphysics: Towards a Theory of Scale Relativity. World Scientific, 1993.
10. Nottale L. Scale relativity and fractal space-time: theory and applications. 2008, arXiv: physics.gen-ph/0812.3857.
11. Rubčić A. and Rubčić, J. Planetary orbits on solar and extrasolar systems. *Fizika A*, 2010, v. 19 (3), 133–144
12. Pintr P., Peřinová V., and Lukš A. Allowed planetary orbits in the solar system. *Chaos, Solitons & Fractals*, 2008, v. 36, 1273–1282.

13. Jaki S. L. The Early History of the Titius-Bode Law. *American Journal of Physics*, 1972, v. 40 (7), 1014–1023
14. Wells D. R. Was the Titius-Bode series dictated by the minimum energy states of the generic solar plasma? *IEEE Transactions on Plasma Science*, 1990, v. 19 (1), 73–76.
15. Wells D. R. Titius-Bode and the helicity connection: a quantized field theory of protostar formation. *IEEE Transactions on Plasma Science*, 1989, v. 14 (6), 865–873.
16. Wells D. R. Quantization effects in the plasma universe. *IEEE Transactions on Plasma Science*, 1989, v. 17 (2), 270–281.
17. Wells D. R. Unification of gravitational, electrical, and strong forces by a virtual plasma theory. *IEEE Transactions on Plasma Science*, 1992, v. 20 (6), 939–943.
18. Surdin M. The origin of the solar system and stochastic electrodynamics. *Il Nuovo Cimento C*, 1980, v. 3 (6), 626–634.
19. Graner F. and Dubrulle B. Titius-Bode laws in the solar system. I. Scale invariance explains everything. *Astronomy and Astrophysics*, 1994, v. 282, 262–268.
20. Graner F. and Dubrulle B. Titius-Bode laws in the solar system. II. Build your own law from disk models. *Astronomy and Astrophysics*, 1994, v. 282, 269–276.
21. Nie Q. Comprehensive research on the origin of the solar system structure by quantum-like model. *International Journal of Astronomy and Astrophysics*, 2010, v. 1. 52–61.
22. Livadiotis G., McComas D.J. Evidence of large-scale quantization in space plasmas. *Entropy*, 2013, v. 15, 1118–1134.
23. Kotliarov I. The Titius-Bode law revisited but not revived. 2008, arXiv: physics.space-ph/0806.3532.
24. Poveda A. and Lara P. The exo-planetary system of 55 Cancri and the Titius-Bode law. *Revista Mexicana de Astronomía y Astrofísica*, 2008, v. 44, 243–246.
25. Patterson C. W. Resonance capture and the evolution of the planets. 1987, *Icarus*, v. 70, 319–333.
26. Zoghbi J.-P. A. Quantization of planetary systems and its dependency on stellar rotation. *Publications of the Astronomical Society of Australia*, 2011, v. 28 (3), 177-201.
27. Peřinová V, Lukš A. and Pintr P. Distribution of distances in the solar system. *Chaos, Solitons & Fractals*, 2007, v. 34 (3), 669–676.
28. Torbett M., Greenberg R., Smoluchowski R. Orbital resonances and planetary formation sites. *Icarus*, v. 49 (3), 313–326.
29. Smarandache F. and Christianto V. Schrödinger Equation and the Quantization of Celestial Systems. *Progress in Physics*, 2006, v. 2, 63–67.
30. Potter F. Multi-planet exosystems all obey orbital angular momentum quantization per unit mass predicted by quantum celestial mechanics (QCM). *Progress in Physics*, 2013, v. 3, 60–61.
31. Müller H. Fractal scaling models of natural oscillations in chain systems and the mass distribution of the celestial bodies in the solar system. *Progress in Physics*, 2010, v. 1, 62–66.
32. Nottale L. Scale-relativity and quantization of the universe. I. Theoretical framework. *Astronomy and Astrophysics*, 1997, v. 327, 867–889.
33. Hermann R., Schumacher G., and Guyard R. Scale relativity and quantization of the solar system. *Astronomy and Astrophysics*, 1998, v. 335, 281–286.
34. Cresson J. The stochastisation hypothesis and the spacing of planetary systems. *Journal of Mathematical Physics*, 2011, v. 52 (11), 113502.
35. Laskar J. On the spacing of planetary systems. *Physical Review Letters*, 2000, v. 84 (15), 3240–3243.
36. Hernández-Mena C. and Benet L. Statistics and universality in simplified models of planetary formation. *Monthly Notes of the Royal Astronomical Society*, 2010, v. 412 (1), 95–106.
37. Rafie F. Quasi-quantization of the orbits in the Solar System. *Astronomical and Astrophysical Transactions*, 2005, v. 24 (2), 81–92.
38. Giné J. On the origin of the gravitational quantization: The Titius–Bode law. *Chaos, Solitons & Fractals*, 2007, v. 32 (2), 362–369.
39. De Neto M., Maia L. A., Carneiro S. An alternative theoretical approach to describe planetary systems through a Schrödinger-type diffusion equation. *Chaos, Solitons & Fractals*, 2004, v. 21 (1), 21–28.
40. Scardigli F. A quantum-like description of the planetary systems. *Foundations of Physics*, 2007, v. 37, 1278–1295.
41. De Neto M. Pythagoras' celestial spheres in the context of a simple model for quantization of planetary orbits. *Chaos, Solitons & Fractals*, 2006, v. 30 (2), 399–406.
42. Moyassari P., and Jalalzadeh S. Weyl geometry approach to describe planetary systems. 2004, arXiv: gr-qc/0410073.
43. Chang Y.-F. Quantized phenomena in astronomy and astronomic quantum theory. *International Journal of Sciences*, 2013, v. 2, 58–73.
44. Minton D. A. and Malhotra R. A record of planet migration in the main asteroid belt. *Science*, 2009, v. 457, 1109–1111.
45. Rein H. Period ratios in multiplanetary systems discovered by Kepler are consistent with planet migration. *Monthly Notices Letters of the Royal Astronomical Society*, 2012, v. 427 (1), L21–L24.
46. Fernández J. A., and W.-H. Ip. Some dynamical aspects of the accretion of Uranus and Neptune - The exchange of orbital angular momentum with planetesimals. *Icarus*, 1984, v. 58, 109–120.
47. Ipatov S. I. Migration of bodies in the accretion of planets. *Solar System Research*, 1993, v. 27, 65–79.
48. Lin D. N. C., Bodenheimer P., and Richardson D.C. Orbital migration of the planetary companion of 51 Pegasi to its present location. *Nature*, 1996, v. 380, 606–607.

SPECIAL REPORT

Forty Lines of Evidence for Condensed Matter — The Sun on Trial: Liquid Metallic Hydrogen as a Solar Building Block

Pierre-Marie Robitaille

Department of Radiology, The Ohio State University, 395 W. 12th Ave, Columbus, Ohio 43210, USA.
robitaille.1@osu.edu

Our Sun has confronted humanity with overwhelming evidence that it is comprised of condensed matter. Dismissing this reality, the standard solar models continue to be anchored on the gaseous plasma. In large measure, the endurance of these theories can be attributed to 1) the mathematical elegance of the equations for the gaseous state, 2) the apparent success of the mass-luminosity relationship, and 3) the long-lasting influence of leading proponents of these models. Unfortunately, no direct physical finding supports the notion that the solar body is gaseous. Without exception, all observations are most easily explained by recognizing that the Sun is primarily comprised of condensed matter. However, when a physical characteristic points to condensed matter, *a postori* arguments are invoked to account for the behavior using the gaseous state. In isolation, many of these treatments appear plausible. As a result, the gaseous models continue to be accepted. There seems to be an overarching belief in solar science that the problems with the gaseous models are few and inconsequential. In reality, they are numerous and, while often subtle, they are sometimes daunting. The gaseous equations of state have introduced far more dilemmas than they have solved. Many of the conclusions derived from these approaches are likely to have led solar physics down unproductive avenues, as deductions have been accepted which bear little or no relationship to the actual nature of the Sun. It could be argued that, for more than 100 years, the gaseous models have prevented mankind from making real progress relative to understanding the Sun and the universe. Hence, the Sun is now placed on trial. Forty lines of evidence will be presented that the solar body is comprised of, and surrounded by, condensed matter. These 'proofs' can be divided into seven broad categories: 1) Planckian, 2) spectroscopic, 3) structural, 4) dynamic, 5) helioseismic, 6) elemental, and 7) earthly. Collectively, these lines of evidence provide a systematic challenge to the gaseous models of the Sun and expose the many hurdles faced by modern approaches. Observational astronomy and laboratory physics have remained unable to properly justify claims that the solar body must be gaseous. At the same time, clear signs of condensed matter interspersed with gaseous plasma in the chromosphere and corona have been regrettably dismissed. As such, it is hoped that this exposition will serve as an invitation to consider condensed matter, especially metallic hydrogen, when pondering the phase of the Sun.

The Sun is a world so different from our own . . . However [relative to understanding its structure], *one must not lose heart; over the past few years science has made a lot of progress, and those who come after us will not fail to make even more.*

Father Angelo Secchi, S.J., 1875 [1, p. 300, V. I][*]

1 Introduction

A long time ago, men like Gustav Kirchhoff, Johann Zöllner, William Thomson (Lord Kelvin), and James Jeans viewed the photosphere (or the solar body) as existing in the liquid state [2, 3]. Despite their stature, scientists, since the days of Herbert Spencer and Angelo Secchi, slowly drifted towards the concept that the Sun was a ball of gas surrounded by condensed matter [2, 3].[†]

Others, of equal or greater prominence, including August Ritter, Jonathan Lane, Franz Schuster, Karl Schwarzschild, Arthur Eddington, Subrahmanyan Chandrashekhar, and John Bahcall, would have their chance to speak [2, 3]. The Sun became a fully gaseous plasma.

As a consequence, the gaseous Sun has imbedded itself at the very foundation of astronomy. Few would dispute that

[*]Translations from French were executed by the author.

[†]In the mid-1800s, five great pillars had given birth to the gaseous Sun: 1) Laplace's Nebular Hypothesis, 2) Helmholtz' contraction theory, 3) Cagniard de la Tour's critical phenomena and Andrew's critical temperatures, 4) Kirchhoff's formulation of his law of thermal emission, and 5) the discovery of pressure broadening in gases. Each of these has previously been addressed in detail [2].

the Sun is a gas and that our understanding of all other stars and the entire universe, is inherently linked to this reality. Therefore, any endeavor to touch the phase of the Sun must be viewed as an attempt to reformulate all of astronomy.

Yet, when astrophysics remained a young science, observational astronomers, such as James Keeler, Edwin Frost, and Charles Abbot [4], objected to the theoretical basis for a gaseous Sun. August Schmidt was the first to mathematically dismiss the solar surface as illusion. Speaking of him, Charles Abbot, the director of the Smithsonian Observatory would write, *"Schmidt's views have obtained considerable acceptance, but not from observers of solar phenomena"* [5, p. 232]. In 1913, Charles Maunder made the point even more forcefully, *"But under ordinary conditions, we do not see the chromosphere itself, but look down through it on the photosphere, or general radiating surface. This, to the eye, certainly looks like a definite shell, but some theorists have been so impressed with the difficulty of conceiving that a gaseous body like the Sun could, under the conditions of such stupendous temperatures as there exist, have any defined limit at all, that they deny that what we see on the Sun is a real boundary, and argue that it only appears so to us through the effects of the anomalous refraction or dispersion of light. Such theories introduce difficulties greater and more numerous than those that they clear away, and they are not generally accepted by the practical observers of the Sun"* [6, p. 28]. Alfred Fowler, the first Secretary of the International Astronomical Union, shared these views, *"The photosphere is thus regarded as an optical illusion, and remarkable consequences in relation to spots and other phenomena are involved. The hypothesis appears to take no account of absorption, and, while of a certain mathematical interest, it seems to have but little application to the actual Sun"* [7].

With time, however, the voices of the observational astronomers were silenced by the power and elegance of the mathematical arguments [2, 3]. Those who could not follow sophisticated theory could no longer become professional astronomers. At Cambridge, the Mathematical Tripos became and remained an accepted path to a Ph.D. degree in astronomy [8]. Theory [9–14],* rather than observation, came to dictate the phase of the Sun and all solar phenomena were explained in terms of a gaseous entity.

As gases are unable to support structure, additional means were adopted to explain solar observations. Magnetic fields became the solution to every puzzle [12], even though gases are incapable of their generation.† Over time, theoretical approaches claimed one victory after the next, until it seemed as if the Standard Solar Models [11, 13, 14] were unshakable. Gases were inappropriately endowed with all of the properties of condensed matter.

In reality, a closer examination would have revealed that many theoretical achievements were inapplicable. Some of the difficulties stemmed from improper experimental conclusions. The universality of several laws [15–20], on which the entire solar framework rested [9, p. 27–58], was the product of faulty assumptions [21–24]. These errors were introduced when theoretical physics remained in its infancy. But now, they were governed by other branches of physics (i.e. blackbody radiation and condensed matter physics [15–20, 25]), not by astronomy. The most pressing problems were never properly solved by the physics community [21–24].

Solar theory was replete with oversights and invalid assumptions, but the shortcomings would be extremely difficult to detect. Problems which were 'solved 100 years ago' still lurked in the background [19,20]. Too much forward progress was desired with too little attention paid to the road traveled. Most viewed that only a few minor problems remained with gaseous equations of state [13,14]. Evidence that the Sun was not a gas was dismissed with complex schemes often requiring the suspension of objectivity.

Nonetheless, many lines of evidence had revealed that the body of the Sun must be comprised of condensed matter (see Table I). Slowly, arguments initially advanced by men like Gustav Kirchhoff [26] and James Jeans [27, 28] began to reemerge. Moreover, they were joined by an arsenal of new observations. Today, at least forty proofs can be found disputing the gaseous nature of the Sun. There are surely more to be discovered.‡ Conversely, not one direct proof exists that the body of the Sun must be considered a gaseous plasma.

It is clear that the lines of evidence for condensed matter which are contained herein§ are worthy of a cohesive discussion. For the purpose of this presentation, they are subdivided and reorganized into seven broad categories: 1) Planckian, 2) spectroscopic, 3) structural, 4) dynamic, 5) helioseismic, 6) elemental, and 7) earthly. Each proof will be discussed relative to the liquid metallic hydrogen (LMH) model [36, 39, 47, 48] wherein condensed hydrogen, pressurized in the solar interior, assumes a graphite-like lattice on the photosphere [39, 40, 45, 48], a more metallic nature in sunspots and faculae [40,45,52], a diffuse presence in a somewhat cool

*Eddington's mass-luminosity relationship [9, p. 145–179] stands as one of the great triumphs of the gaseous models. Today, this finding is well established in observational astronomy and Eddington's derivation is worthy of a detailed treatment. Due to space limitations, the topic will not be addressed herein. Suffice it to state that Eddington's derivation was dependent on the validity of Kirchhoff's law and no effort has been made to account for the relationship if the stars were made of condensed matter. At the same time, it must be noted that through the mass-luminosity relationship, an observation linked to distant objects, came to dictate the phase of the Sun. The relationship is not contingent on the behavior of the Sun itself, although the latter does lie on the main sequence of the stars.

†Magnetic fields are the product of underlying microscopic structure in condensed matter. As such, whenever a magnetic field is generated on Earth, condensed matter must be involved, either to directly generate it, or to cause the ordered flow of charge.

‡Solar astronomers, upon further consideration, will recognize that their own subject areas might also provide additional lines of evidence. With time, these complimentary proofs will eventually surface.

§The author presents a complete list of his relevant works [2–4, 29–62] in order to facilitate the study of these problems.

Table 1: Forty Lines of Evidence for Condensed Matter — The Sun on Trial.

corona [57, 58, 60], and a solid character in the core [50].*

Of these lines of evidence, the thermal proofs will always remain central to understanding the condensed nature of solar material. They are tied to the most important questions relative to light emission [15–20] and have the ability to directly link physical observation to the presence of a vibrational lattice, a key aspect of all matter in the condensed phase [21–24]. Hence, the discussion begins with the thermal lines of evidence, as inherently related to blackbody radiation [15–25, 63] and to the earliest scientific history of the Sun [2, 3].

2 Planckian (or Thermal) Lines of Evidence

The Sun emits a spectrum in the visible and infrared region of the electromagnetic spectrum (see Fig. 1) whose detailed analysis provides a total of eight lines of evidence relative to the presence of condensed matter.† For gaseous models, solar emission must be explained using the most complex of schemes, resting both on the validity of Kirchhoff's law of thermal emission [15, 16] and on the *'solar opacity problem'* [42].

Agassi reminds us that *"Browsing through the literature, one may find an occasional use of Kirchhoff's law in some*

*The model adopts a liquid state for the surface of the Sun, as this is in keeping with macroscopic observations. However, an extended structural lattice, not simply a random assembly of degenerate atoms, is required, as demonstrated in §2. Of course, on the scale of solar dimensions, even a material with the rigidity of a solid on Earth (i.e. with a high elastic modulus), might well appear and behave macroscopically as a liquid on the photosphere.

†These proofs require the longest descriptions, as they touch many concepts in physics. Since they deal with thermal phenomena, they can also be referred to as the *'Planckian'* lines of evidence, in recognition of Max Planck's contribution to this area of physics [19, 20]. Beyond physics, Max Planck's philosophical writings (see references in [64]) and personal conduct [65], despite the evil of his times, have much to offer to modern society.

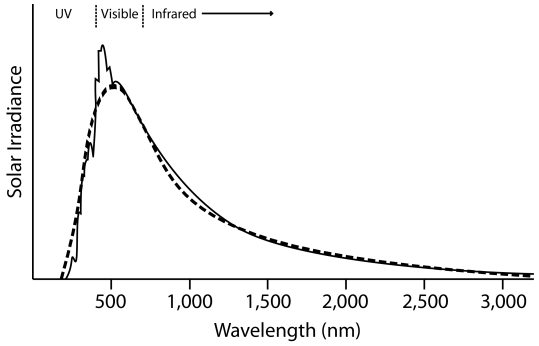

Fig. 1: Schematic representation of the visible spectrum of the Sun (adapted from Fig. 1–3 in [66]). To a first approximation, the solar spectrum is very nearly identical to that of a blackbody with a temperature of ∼5,800 K (dashed line).

experimental physics, but the only place where it is treated at all seriously today is in the astrophysical literature" [63]. In reality, it would not be an overstatement to argue that Kirchhoff's law [15, 16] constitutes the very core of accepted solar theory. Any problems with its formulation would send shock waves not only throughout stellar astrophysics, but to every corner of modern astronomy. Hence, the discussion with respect to the thermal lines of evidence commences with a review of Kirchhoff's law [15, 16] and of blackbody radiation [17–25]. This will be followed by an overview of these principles, as applied to the Sun and the resulting solar opacity problem [42].

2.1 Blackbody Radiation and Kirchhoff's Law

The author has previously stated that, *"Kirchhoff's law is one of the simplest and most misunderstood in thermodynamics"* [24].* Formulated in 1860 [15,16], the law was advanced to account for the light emitted from objects in response to changes in temperature. Typically, in the mid-1800s, the objects were black, as they were covered with soot, or black paint, for best experimental results [21, 23, 24]. Thus, this field of research became known as the study of *'blackbody radiation'* [21, 23, 24]. Kirchhoff attempted to synthesize an overarching law into this area of physics in order to bring a certain unification to laboratory findings. At the time, physics was in its infancy and theorists hoped to formulate laws with *'universal'* consequences. Such was Kirchhoff's goal when his law of thermal emission was devised.

The heart of Kirchhoff's law states that, *"If a space be entirely surrounded by bodies of the same temperature, so that no rays can penetrate through them, every pencil in the interior of the space must be so constituted, in regard to its quality and intensity, as if it had proceeded from a perfectly black body of the same temperature, and must therefore be independent of the form and nature of the bodies, being determined by the temperature alone . . . In the interior therefore of an opake red-hot body of any temperature, the illumination is always the same, whatever be the constitution of the body in other respects"* [16, §16].†

Blackbody radiation was governed strictly by the temperature and the frequency of interest. *The nature of the walls was irrelevant.* Kirchhoff introduced the idea that blackbody radiation somehow possessed a *'universal'* significance and was a property of all cavities [15, 16].

Eventually, Max Planck [19, 20] provided a mathematical form for the spectral shape of blackbody emission sought by Kirchhoff [15, 16]. Kirchhoff's law became ingrained in Planck's formulation [20, §24–§62]. By extension, it also became an integral part of the laws of Wien [17] and Stefan [18], as these could be simply derived from Planck's equation [20, §31–§60]. In turn, the laws of radiation, came to form the very foundation of the gaseous models (see e.g. [9, p. 27–58]).

Since blackbody radiation was thought to be of a *'universal'* nature and *independent of the nature of the walls*, Max Planck, was never able to link his equation to a direct physical cause [21, 23, 24].‡ He spoke of any such attempt as a *'hopeless undertaking'* [20, §41]. In this respect, blackbody radiation became unique in physics. Planck's equation was not linked to anything in the material world, as Kirchhoff's law [15, 16] had dictated that the process was detached from physical causality [20, 21].

With his law, Gustav Kirchhoff was informing the physics community that the light emitted by an object will always correspond to the same *'universal'* spectrum at a given temperature, provided that the object be enclosed and the entire system remain at thermal equilibrium. Any enclosure contained the same blackbody radiation. The nature of the enclosure was not relevant to the solution, given that it was truly opaque. Perfectly reflecting enclosures, such as those made from silver, should function as well as perfectly absorbing enclosures made from graphite or coated with carbon black.

In reality, Kirchhoff erred in believing that the nature of

*A detailed series of publications related to the analysis of Kirchhoff's law has previously appeared. These can be consulted by those who seek a more extensive discussion of the subject matter (see [21–24]).

†Note how this last sentence immediately implied that, if the solar interior could be viewed as enclosed, then the radiation existing within it must be of the same form (intensity versus frequency) as that emitted by a blackbody at the temperature in question.

‡In processes where light is emitted, there are five aspects to consider: 1) the physical setting, 2) separate energy levels created in this setting, 3) a transition species which will make use of these energy levels, 4) the production of a photon, and 5) an equation. For instance, for Lyman-α radiation these correspond to 1) the hydrogen atom, 2) the two electronic orbitals involved in the transition — principle quantum numbers N=2 and N=1, 3) the electron as the transition species, 4) the Lyman-α emission at 1216Å, and 5) the Rydberg formula. Alternatively, in speaking of the proton nuclear magnetic resonance line from water, these correspond to 1) the hydrogen atoms of the water molecules placed in a magnetic field, 2) the hydrogen nuclear spin up or spin down states, 3) the hydrogen nuclear spin as a transition species, 4) the hydrogen line at 4.85 ppm, and 5) the Larmor equation. Analogous entries can be made for any spectroscopic process in physics, with the exception of blackbody radiation. In that case, only the 4th and 5th entries are known: 4) the nature of the light and 5) Planck's equation [21].

the enclosure did not matter [21–24]. Perfectly reflecting enclosures manifest the radiation of the objects they contain, not blackbody radiation (see [22] for a proof). To argue otherwise constitutes a violation of the First Law of Thermodynamics. Furthermore, if Kirchhoff's law was correct, any enclosed material could serve as an experimental blackbody. But, laboratory blackbodies are known to be extremely complex devices, typically involving the use of specialized *'nearly perfectly absorbing'* materials over the frequencies of interest.*

Max Planck believed that *"...in a vacuum bounded by totally reflecting walls any state of radiation may persist"* [20, §61]. In itself, this was contrary to what Kirchhoff had stated, as noted above, *"...In the interior therefore of an opake red-hot body of any temperature, the illumination is always the same, whatever be the constitution of the body in other respects"* [16, §16]. Throughout his text on thermal radiation [20], Max Planck repeatedly introduces a *'small carbon particle'* to ensure that the radiation he was treating was truly black [21, 23]. He viewed the particle as a *catalyst* and believed that it simply accelerated the move towards black radiation. In reality, he had introduced a perfect absorber/emitter and thereby filled the cavity with the radiation desired (see [22] for a proof). If Kirchhoff's law was correct, this should not be necessary. The carbon particle was much more than a simple catalyst [21, 23].

Another repercussion to Kirchhoff's statement was the belief that objects could radiate internally. In fact, Planck would use this approach in attempting to derive Kirchhoff's law (see [20, p. 1–45]).† Yet, conduction and/or convection properly govern heat transfer within objects, not internal radiation. Thermal radiation constitutes an attempt to achieve equilibrium with the outside world.

The idea that all opaque enclosures contain blackbody radiation was demonstrably false in the laboratory and Kirchhoff's law of thermal emission, invalid [21–24].‡ Rather, the best that could be said was that, at thermal equilibrium and in the absence of conduction or convection, the absorption of radiation by an object was equal to its emission. This was properly formulated by Balfour Stewart in 1858, one year before Kirchhoff developed his own law [22, 25].

The universality which Kirchhoff sought was not present. Regrettably, Max Planck had embraced this concept and, as a direct consequence, blackbody radiation was never linked to a direct physical cause. Tragically, the astrophysical community would come to believe that blackbody radiation could be produced without the presence of condensed matter. Upon this *ex nihilo* generation, it built the foundations of a gaseous Sun [9, p. 27–58] and the framework of the universe.

2.2 Kirchhoff's Law, Solar Opacity, and the Gaseous Models of the Sun

Given thermal equilibrium, Kirchhoff's belief that all opaque enclosures contained blackbody radiation had profound consequences for astronomy. If the Sun was considered to be an enclosure operating under thermal equilibrium, then by Kirchhoff's law, it was filled with blackbody radiation (e.g. [9, p. 27–58]). Nothing was required to produce the radiation, other than adherence to Kirchhoff's condition. Even so, use of the laws of thermal emission [15–20] explicitly required the presence of thermal equilibrium in the subject of interest (i.e. conduction and convection must not be present [21–24]).

As for the Sun, it operates far out of equilibrium by every measure, emitting a large amount of radiation, but absorbing essentially none. Furthermore, it sustains clear differential convection currents on its surface, as reported long ago by Carrington [67, 68]. Consequently, how could the proponents of the gaseous models justify the use of the laws of thermal emission to treat the interior of the Sun [9, 13, 14]? How could an object like the Sun be considered enclosed?

Arthur Eddington viewed the Sun as filled with radiation which was essentially black. For him, the Sun acted like a slowly leaking sieve [9, p. 18]. In speaking of the application of Stefan's law [18] to the solar interior, Eddington argued, *"To a very high degree of approximation the last two results are immediately applicable to the interior of a star. It is true that the radiation is not in an ideal enclosure with opaque walls at constant temperature; but the stellar conditions approach the ideal far more closely than any laboratory experiments can do"* [9, p. 99–100]. He justified these statements based on the *very opaque* nature of stellar material which he inferred by considering a distant star, Capella [9, p. 100].

Stefan's law codified a fourth power dependence on temperature (T^4) [18]. At the same time, the gaseous Sun was thought to sustain a core temperature of roughly 1.6×10^7 K [13, p. 9] while displaying an apparent surface temperature of only 6,000 K. Therefore, application of Stefan's law [18] to imaginary concentric spheres [13, p. 2] located in the interior of the Sun would result in a great deal more photons produced in the core than ever emitted by its surface. Through the application of such logic, the Sun could be viewed as a slowly leaking sieve and essentially perfectly enclosed. Eddington inferred that the opacity, or ability to absorb a photon, within the Sun was extremely elevated. Under these circumstances,

*For an extensive list of references on laboratory blackbodies and the materials used in their preparation, see [23].

†In his derivation, Planck did not permit his volume-elements to reflect light [20, p. 1–45]. As a result, all these elements became perfectly absorbing and he was able to obtain Kirchhoff's law. However, had he properly included reflection, he would have convinced himself that Kirchhoff's law was invalid (see [21–24] for a complete discussion).

‡One cannot expect scientists to revisit the validity of every law upon which they shall base their work. As such, if 20th century astronomers committed a misstep in applying Kirchhoff's law to the Sun, it is not at all clear how this could have been prevented. Indeed, when the author was first considering these problems, he actually believed that Kirchhoff's law was valid (i.e. [29]), but that the Sun simply failed to meet the requirements set forth by enclosure. It was only later, following an extensive review of blackbody radiation [21–24], that he came to realize that there was an error in the law itself.

light produced in the solar interior could not travel very far before being absorbed (see [9, p. 100] and [14, p. 185–232]).*

Arthur Milne argued that the interior of a star could be viewed as being in *local thermal equilibrium*, thereby insisting that Kirchhoff's law could be applied within the Sun. Speaking of the solar interior, he stated, *"If the atoms are sufficiently battered about by colliding with one another, they assume a state (distribution of stationary states) characteristic of thermodynamic equilibrium at temperature T"* [69, p. 81–83]. Unfortunately, these words describe the conditions required for the onset of conduction [70]. Thermal equilibrium could never exist at the center of the Sun, as the setting prevailing at the core would facilitate a non-radiative process [21–24].†

Max Planck has clearly stated that thermal equilibrium can only exist in the absence of all conduction, *"Now the condition of thermodynamic equilibrium requires that the temperature shall be everywhere the same and shall not vary with time . . . For the heat of a body depends only on the heat radiation, since on account of the uniformity in temperature, no conduction of heat takes place."* [20, §24]. That is why he insisted that the walls of the enclosure be rigid (e.g. [20, §24–25]), as no energy must be carried away through the action of the momentum transfer which accompanies collisions. Accordingly, Milne's arguments, though they rest at the heart of the gaseous solar models, are fallacious. It is inappropriate to apply Stefan's law to the interior of the Sun, as conductive forces violate the conditions for enclosure and the requirements for purely radiative heat transfer.‡

In his treatise on heat radiation, Planck warned against applying the laws of thermal emission directly to the Sun, *"Now the apparent temperature of the Sun is obviously nothing but the temperature of the solar rays, depending entirely on the nature of the rays, and hence a property of the rays and not a property of the Sun itself. Therefore it would be not only more convenient, but also more correct, to apply this notation directly, instead of speaking of a fictitious temperature of the Sun, which can be made to have meaning only by the introduction of an assumption that does not hold in reality"* [20, §51]. Planck must have recognized that the Sun possessed convection currents on its surface [41], as Carrington's discovery [67] would have been well-established throughout scientifically educated society.

To further complicate matters, astrophysics must create sufficient opacity in the Sun. Opacity acts to contain and shift the internal radiation essential to the gaseous models. It has been said that absorption of radiation in the solar interior takes place through the summation of innumerable processes (including bound-bound, bound-free, free-free, and scattering reactions [14, p. 185–232]). Such a hypothesis constitutes the *'stellar opacity problem'*.§ The blackbody spectrum which could be produced in the laboratory using simple materials like graphite, soot, or metal-blacks [21–24], at once required the summation of a large set of processes which were not known to contribute to the production of the blackbody spectrum on Earth [41, 42]. The central problem for gas models is not that the Sun sustains clear convection at the level of the photosphere, nor that inferred conduction exists at its core. Rather, it was that Kirchhoff's law was not valid and that Planck's equation had not been linked to the physical world [21–24]. The laws of thermal emission could not be applied to the Sun. It was not reasonable to account for the production of a blackbody spectrum using opacity calculations which depended on processes unrelated to thermal emission [42]. The production of blackbody radiation required much more than imaginary enclosures. It required the presence of nearly perfectly absorbing condensed matter, as well-demonstrated by all laboratory experiments over the course of more than 200 years (see [21–24] and references therein).

2.3 The Eight Planckian Lines of Evidence

The eight Planckian (or thermal) lines of evidence, on their own, provide sufficient proof that the Sun is comprised of condensed matter. Each of these proofs includes two components 1) a discussion of some aspect of thermal radiation, and 2) the associated structural implications. It has been well-established in experimental physics that the thermal emissivity of a material is directly linked to its structure [71]. Furthermore, condensed matter is known to possess varying directional emissivities which play a key role in understanding the structures associated with the Sun, including the degree to which one might infer that they are metallic [66, 72, 73].

2.3.1 Solar Spectrum #1

The blackbody lineshape of the solar spectrum (see Fig. 1) has been known since the days of Samuel Langley (see [74, Plate 12 and 21] and [75, Plate IV]).¶ Still, though astrophysics

*Eddington concluded that *"the stars on the main series possess nearly the same internal temperature distribution"* and inferred core temperatures in the millions of degrees [9, p. 177–178]. Given his belief that the laws of thermal emission [15–20] could be applied to the core of the stars, the temperatures he inferred would result in the production of photons with X-ray energies. Over thousands of years, these photons would slowly work their way out to escape at the photosphere. But as they traveled to the surface, they would slowly lose energy and become shifted to ever lower frequencies. Finally, upon reaching the surface, they would emit in the visible region of the electromagnetic spectrum. To accomplish the feat, the gas models required that perfect and gradual changes in opacity enabled a blackbody spectrum produced at X-ray frequencies to be slowly converted to one existing in white light. The issue has previously been addressed by the author [3, 36, 42] and provides an example where accepted science required the suspension of disbelief.

†The density at the center of the Sun is believed to approach 150 g/cm^3 [14, p. 483], a value compatible with conductive solids on Earth.

‡The Sun is known to possess powerful magnetic fields and a solar dynamo. Their existence strongly argues for conduction within condensed matter (see [35, 39] and §5.3).

§The author has previously addressed the stellar opacity problem [42].

¶The first Planckian proof [45] was initially treated in [29, 35, 36, 42, 43].

has tried to explain the production of this light for nearly 150 years [2, 3], little real progress has been made in this direction. As demonstrated in Section 2.2, the gaseous models fail to properly account for the occurrence of the solar spectrum. Gases are unable to emit a continuous spectrum. Rather, they emit in bands (see [21,70] and references therein). Even when pressure broadened, these bands cannot produce the blackbody lineshape. Moreover, when gases are heated, their emissivity can actually drop [21,70], in direct contradiction of Stefan's law [18]. Under these circumstances, the answer cannot be found in the gaseous state. One must turn to condensed matter.

Throughout history, the production of a blackbody spectrum [21, 23, 24] has been facilitated by the use of graphite [76–84] or soot. For this reason, even after the formulation of Kirchhoff's law, astronomers envisioned that graphite particles floated on the surface of the Sun [2, 3]. Hastings recognized that the solar surface was too hot to permit the existence of carbon in the condensed state [85]. He noted that *"Granting this, we perceive that the photosphere contains solid or liquid particles hotter than carbon vapor, and consequently not carbon"* [85]. As a result, in 1881, he suggested that *"... the substance in question, so far as we know it, has properties similar to those of the carbon group"* [85]. Hastings wanted something which had the physical characteristics of graphite, especially related to emissivity. Yet, the only aspect of graphite which could contribute to its emissive characteristics was its lattice structure. He was indirectly searching for a material which might share the lattice arrangement known to exist in graphite (see Fig. 2), but which might likewise be reasonably expected to exist on the surface of the Sun.

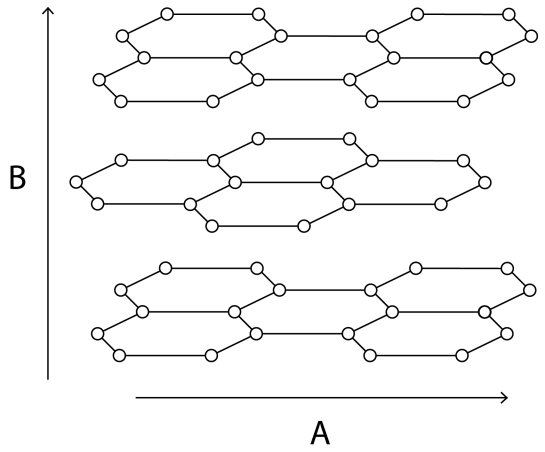

Fig. 2: Schematic representation of the layered hexagonal lattice found in graphite (adapted from Fig. 1 in [48]).

Eventually, Cecilia Payne determined that the stars were largely made of hydrogen [86] and Henry Norris Russell [87] extended the conclusion to the Sun.* Whatever was responsible for the thermal spectrum had to be composed of hydrogen.

Then, in 1935, a seminal work appeared which had the potential to completely alter our understanding of the stars [36, 39]. Eugene Wigner (Nobel Prize, Physics, 1963) and H.B. Huntington [88], proposed that at sufficient pressures, hydrogen could become metallic. More importantly, they would make a direct link between the structure of metallic hydrogen and that of graphite itself, *"The objection comes up naturally that we have calculated the energy of a body-centered metallic lattice only, and that another metallic lattice may be much more stable. We feel that the objection is justified. Of course it is not to be expected that another simple lattice, like the face-centered one, have a much lower energy, — the energy differences between forms are always very small. It is possible, however, that a layer-like lattice has a much greater heat of formation, and is obtainable under high pressure. This is suggested by the fact that in most cases of Table I of allotropic modifications, one of the lattices is layer-like*[1]*..."* [88]. The footnote in the text began, *"Diamond is a valence lattice, but graphite is a layer lattice ..."* [88].

With time, Brovman et al. [89] would propose that metallic hydrogen might be metastable. Like diamonds, it would require elevated pressures for formation, but remain stable at low pressures once synthesized. Neil Ashcroft and his group hypothesized that metallic hydrogen might be metastable between its solid and liquid forms [90, 91].

Metallic hydrogen remains elusive in our laboratories (see [39, 92] for recent reviews). Nonetheless, this has not prevented astrophysics from invoking its existence within brown dwarfs and giant planets [93–95], or even in neutron stars [96]. In fact, based on expected densities, temperatures, and elemental abundances obtained using the gaseous models for the solar core, metallic hydrogen has been said to exist at the center of the Sun [97–99].†

In previous astrophysical studies [93–99], thermal emission has not guided the selection of the form which metallic hydrogen would adopt. As a result, they have sidestepped the layered graphite-like structure first suggested by Wigner and Huntington [88]. Nonetheless, it seems clear that metallic hydrogen, based on the inferred solar abundance of hydrogen [86, 87] and extensive theoretical support (see [39, 92] for

*See [47] for a detailed discussion on the composition of the Sun.

†Setsuo Ichimaru was primarily concerned with nuclear reactions in high density plasmas [97–99]. His work on the solar core is based on assumptions for the composition of the solar interior [97, p. 2] which are derived from the gaseous models, *"In the Sun ... the mass density and the temperature are estimated to be 156 g/cm^3 and 1.55x10^7, respectively. The mass fraction of hydrogen near the core is said to be 0.36 and thus the mass density of metallic hydrogen there is 56.2 g/cm^3"* [98, p. 2660]. Ichimaru places specific emphasis on the One-Component Plasma (OCP) [97, pp. 103 & 209]. He assumed that the lattice points were those of a body-centered cubic [97]. The body-centered cubic is a solid structure. Its existence within the Sun had not been justified beyond inferred densities. Ichimaru's assumptions would have been easily supported by recent seismological evidence which demonstrates that the solar core experiences solid body rotation (see [50] and §6.5 in this work). His supposition has important consequences for driving nuclear reactions within the Sun (see [44, 48] and §7.1 in this work).

reviews), constitutes an ideal building material for the entire Sun which is appropriate for 21st century thought.

Thus, theoretical condensed matter physics unknowingly provided astronomy with everything needed to explain the origin of the thermal spectrum (see Fig. 1). Payne and Russell had determined that the Sun was composed of hydrogen [86, 87]. Under the enormous pressures which existed in the solar interior, Wigner and Huntington [88] allowed that this hydrogen could be converted to the metallic state and adopt the lattice structure of graphite. Work by Brovman et al. [89] enabled metallic hydrogen, formed under high pressure conditions within the solar interior, to be metastable at the surface. Thermal emission could then result from lattice vibrations [21], occurring within layered metallic hydrogen, much like what occurs with graphite on Earth.

In contrast to the gaseous models, where photons take millions of years to escape from the solar core [9], in a liquid metallic hydrogen (LMH) Sun, light can be instantly produced at the level of the photosphere, using mechanisms identical to those found within graphite. Complex changes in internal solar opacities are not required [42]. The solar spectrum can be explained without recourse to unsuited gases [21, 70], imaginary enclosures [9], dismissal of observed conduction [69] and convection [67, 68], the need for local thermal equilibrium [69], or Kirchhoff's erroneous law [15, 16]. The conjecture that solar thermal emission is produced by hydrogen in the condensed state on the surface of the Sun is simpler than any scheme brought forth by the gaseous models. Furthermore, it unifies our understanding of thermal emission in the stars with that of laboratory models on Earth. But most importantly, it results in the incorporation of a structural lattice directly onto the photosphere, providing thereby a basis upon which every other physical aspect of the Sun can be directly explained — from the presence of a true surface to the nature of all solar structures. Hydrogen's ability to exist as condensed matter within the solar body, photosphere, chromosphere, and corona, appears all but certain. The remainder of this work should help to further cement this conclusion.

Fig. 3: Image of the Sun displaying how the intensity of the disk decreases towards the limb [100]. Note this image was described as follows, *"Sunspot group in context. The diameter of the Sun is 100 times larger than the diameter of the Earth. This image was recorded with our finder telescope at about the same time as the 15 July images and movies. Target: The Sun; Date: 15 Jul 2002"*. It is reproduced herein thanks to the generosity of the Royal Swedish Academy of Sciences (www.solarphysics.kva.se/NatureNov2002/press_images_eng.html — accessed online 9/15/2013). The SST is operated on the island of La Palma by the Institute for Solar Physics of the Royal Swedish Academy of Sciences in the Spanish Observatorio del Roque de los Muchachos of the Instituto de Astrofísica de Canarias.

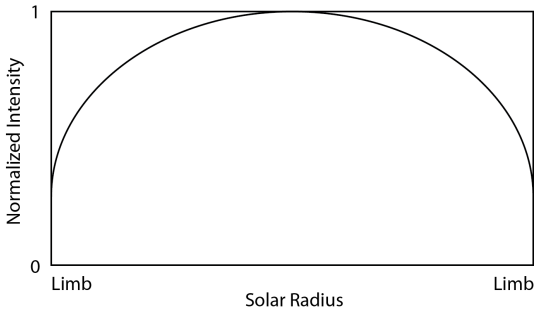

Fig. 4: Schematic representation of the white light intensity variation across the solar disk which is responsible for visible limb darkening. The extent of intensity variation is frequency dependent [101].

2.3.2 Limb Darkening #2

According to Father Angelo Secchi, while Galileo denied the existence of limb darkening (see Figs. 3, 4), the phenomenon had been well established by Lucas Valérius of the Lincei Academy, *"...the image of the Sun is brighter in the center than on the edges."* [1, p. 196, V. I].*

In 1902, Frank Very demonstrated that limb darkening was a frequency dependent phenomenon [101] which he attributed to scattering in the solar atmosphere and reflection with carbon particles.†

Very's study of solar emission [101] eventually led to the *law of darkening* initially developed by Karl Schwarzschild [102], whereby the observed phenomenon could be explained by relying on the assumption that radiative equilibrium existed within the stars. Once again, this was viewed as a great triumph for gaseous models (see [3] for additional details).

Arthur Eddington would come to adopt Milne's treatment [103] of the *law of darkening* [9, p. 320–324]. However, all of these approaches shared a common flaw: they were based on the validity of Kirchhoff's law [15, 16]. Karl Schwarzschild's derivation began with the words, *"If E is the emission of a black body at the temperature of this layer and one assumes that Kirchhoff's law applies, it follows that the layer will radiate the energy Eadh in every direction"* [102, p. 280 — in Meadows].

Beyond the validity of Kirchhoff's law, these derivations sidestepped the reality that clear convection currents existed

*The second Planckian proof [45] was initially treated in [3, 35, 40, 42].

†As nearly perfect absorbers, carbon particles make for poor reflectors.

on the exterior of the Sun [67, 68]. Remarkably, just a few years after publishing his classic derivation of the law of darkening [103], Milne himself argued that local thermal equilibrium did not apply in the outer layers of the stars [69]. Arthur Eddington also recognized that the laws of emission could not be used to treat the photosphere, *"The argument cannot apply to any part of the star which we can see; for the fact that we see it shows that its radiation is not 'enclosed' "* [9, p. 101]. As such, how could Kirchhoff's law be invoked to explain limb darkening?

To further complicate the situation, any explanation of limb darkening for gaseous models would once again resurrect the solar opacity problem [42]. How could the exterior of the Sun generate a perfect blackbody spectrum using an assembly of processes not seen within graphite?

Gas models accounted for limb darkening by insisting that the observer was sampling different depths within the Sun (see Fig. 5). When viewing the center of the disk, our eye was observing radiation originating further in the interior. This radiation was being released from a layer which was at a higher temperature. Hence, by the Wien's law [17] it appeared brighter. As for limb radiation, it was being produced at shallower depths, thereby appearing cooler and darker.

These ideas were reliant on the belief that the surface of the Sun was merely an illusion,* a conjecture which will be refuted in §3.1, §3.2, §3.7, §4.3, §4.5, §5.1, §5.2, §5.5, §5.7, §6.1, §6.2, and §6.3.

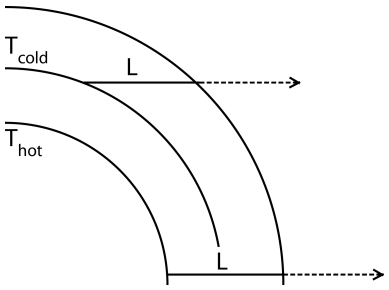

Fig. 5: Schematic representation of how limb darkening is explained in the gas models. When viewing the center of the solar disk, the line of sight travels to a greater depth (L), where it reaches a hotter layer in the solar body. Conversely, when the limb is visualized, the line of sight (L) is restricted to a cooler upper layer. One of the fallacies of this explanation is that the outer layers of the photosphere cannot be considered enclosed (i.e. we can see through them when we visualized the center of the disk). So, photospheric radiation could not be blackbody, even assuming that Kirchhoff's law was valid. Eddington himself had reached this conclusion [9, p. 101].

In the end, the simplest explanation for limb darkening lies in the recognition that directional spectral emissivity occurs naturally within condensed matter [66, 71–73]. Poor conductors tend to have elevated normal emissivities which gradually fall as the angle of observation is decreased (see Fig. 6). This is precisely what is being observed across the solar disk. Good conductors often display lower normal emissivities, which can gradually increase as the angle of observation is decreased, prior to decreasing rapidly as the viewing angle becomes parallel to the surface (see Fig. 6).

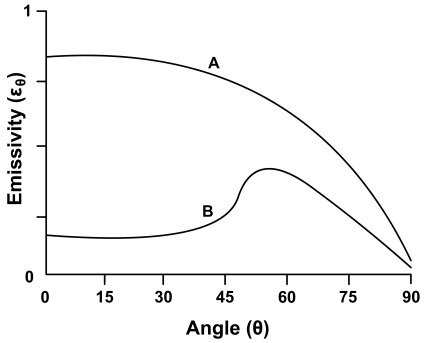

Fig. 6: Schematic representation of directional spectral emissivities for non-conductors (A) and conductors (B). Note that in non-metals, the spectral emissivity decreases monotonically with viewing angle. Conversely, in metals, while the normal emissivity can be substantially reduced, the emissivity can rise with increasing angle before precipitously dropping (adapted from [72]).

Limb darkening revealed that the solar photosphere was condensed, but not highly metallic.† Graphite itself behaves as an excellent emitter, but only a modest conductor. It can be concluded, based on Figs. 4 & 6, that the liquid metallic hydrogen which comprises the solar surface is not highly metallic. The inter-atomic distances in this graphite-like layered material (a Type-I lattice) would be slightly larger than those found in the more metallic sunspots (a Type II lattice), as previously described by the author [35, 39, 40].

2.3.3 Sunspot Emissivity #3

Galileo viewed sunspots (see Fig. 7) as clouds floating very near the solar surface [105].‡ His great detractor, Christoph Scheiner, initially saw them as extrasolar material [2], but eventually became perhaps the first to view them as cavities [1, p. 15, V. I]. This apparent depression of sunspots was confirmed by Alexander Wilson [2] who, in 1774 [106], used precise geometric arguments to establish the effect which now bears his name [1, p. 70–74]. In 1908, George Ellery Hale discovered that sunspots were characterized by intense magnetic fields [107]. This remains one of the most far reaching findings in solar science.

*To this day, astronomy continues to maintain that the Sun's surface is an illusion, as seen in this text produced by the National Solar Observatory, *"The density decreases with distance from the surface until light at last can travel freely and thus gives the illusion of a visible surface"* [104, p. 4].

†As a side note, Frank Very had suggested [101] that the limb darkening of the Sun might be associated with the solar granulations [3, 101]. As will be seen in §2.3.4, the thought was not without merit.

‡The third Planckian proof [45] was initially the 13th line of evidence [35]. It has been presented, in greater detail, within [4, 40, 45].

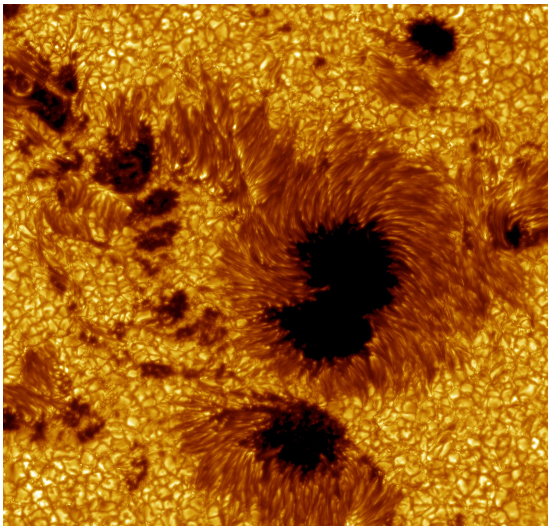

Fig. 7: Part of a sunspot group near the disk center acquired with the Swedish 1-m Solar Telescope [100]. This image has been described as follows by the Institute for Solar Research of the Royal Swedish Academy of Sciences, *"Large field-of-view image of sunspots in Active Region 10030 observed on 15 July 2002. The image has been colored yellow for aesthetic reasons ... Dark penumbral cores — Observations: Göran Scharmer, ISP; Image processing: Mats Löfdahl, ISP; Wavelength: 487.7 nm; Target: AR10030; Date: 15 Jul 2002"*. This image is available for publication thanks to the generosity of the Royal Swedish Academy of Sciences (www.solarphysics.kva.se/NatureNov2002/press_images_eng.html — accessed online 9/15/2013). The SST is operated on the island of La Palma by the Institute for Solar Physics of the Royal Swedish Academy of Sciences in the Spanish Observatorio del Roque de los Muchachos of the Instituto de Astrofísica de Canarias.

In addition to the Wilson effect, sunspot emissivity has been found to drop significantly with increasing magnetic field strength [108, 109]. The magnetic fields within sunspot umbra are known to have a vertical orientation. Their intensity increases in the darkest regions of the umbra (e.g. [110, p. 75] and [111, p. 80]). Sunspot emissivity has also been hypothesized to be directional, with increasing emissivity towards the limb [111, p. 75–77]. In this regard, Samuel Langley had observed, *"With larger images and an improved instrument, I found that, in a complete ring of the solar surface, the photosphere, still brilliant, gave near the limb absolutely less heat than the umbra of the spots"* [112, p. 748]. Edwin Frost echoed Langley, *"A rather surprising result of these observations was that spots are occasionally relatively warmer than the surrounding photosphere"* [113]. Today, the apparent directional changes in the emissivity of sunspots has been dismissed as due to *'stray light'* [111, p. 75–77].

Since a gaseous Sun is devoid of a real surface, the *'Wilson Effect'* cannot be easily explained within these bounds. Once again, optical depth arguments must be made (e.g. see [110, p. 46] and [114, p. 189-190]). In order to account for the emissivity of sunspots, gaseous models propose that magnetic fields prevent the rising of hot gases from the solar interior [104]. Hence, the spot appears cool. But sunspots can possess light bridges (see Secchi's amazing Fig. 33 in [1, p. 69, V. I]). These are characterized by higher emissivities and lower magnetic fields [111, p. 85–86]. The problem for the gaseous models is that light bridges seem to *'float'* above the sunspot. How could these objects be warmer than the material below? Must a mechanism immediately be found to heat light bridges? Sunspots are filled with substructure, including that which arises from Evershed flow. Such substructure is well visible in Fig. 7. However, gases are unable to support structure. How can a gaseous solar model properly account for Evershed flow, while dismissing the surface as an illusion? The problem, of course, remains that all these illusions actually are behaving in systematic fashion (see §5.1). Furthermore, in modern astronomy, the apparent change in sunspot emissivity towards the limb must be dismissed as a *'stray light'* effect. But the most pressing complication lies in the reality that gases are unable to generate powerful magnetic fields (see §5.3). They can respond to fields, but have no inherent mechanism to produce these phenomena. Along these lines, how can magnetic fields be simultaneously produced by gases while at the same time prevent them from rising into the sunspot umbra? On Earth, the production of powerful magnets involves the use of condensed matter and the flow of electrons within conduction bands, not isolated gaseous ions or atoms (see §5.3).

In contrast to the gaseous models, the idea that the Sun is comprised of condensed matter can address all of these complications. The *'Wilson Effect'*, one of the oldest and simplest of solar observations, can continue to be explained without difficulty by using elementary geometry [106], precisely because a true surface can be invoked [45]. The lowered emissivity of sunspot umbra, in association with increased magnetic field strengths, strongly suggests that sunspots are metallic in nature. Langley's observation that sunspots display increased limb emissivity relative to the photosphere can be explained as related to metallic effects.* The increased emissivity and lower magnetic field strength observed within light bridges could be explained by assuming that they, like the photosphere, are endowed with a Type I lattice [35, 39, 40] with lowered metallic properties. Conversely, the decreased normal emissivity of sunspot umbra along with their increased magnetic field strength suggests a more metallic Type II lattice [35, 39, 40] in these structures.

In sunspots, the electrons responsible for generating magnetic fields can be viewed as flowing freely within the conduction bands available in metallic hydrogen. This implies

*This is not to say that stray light cannot present problems. However, these effects should make faculae even less apparent towards the limb, further highlighting the importance of the increase in emissivity which those structures display (see §2.3.5). Definitive answers may come eventually by examining large sunspots.

that the lattice within sunspot umbrae are positioned so that the hexagonal hydrogen planes (see direction A in Fig. 2) are nearly orthogonal to the solar surface (see Fig. 8). In the penumbra, they would be oriented more horizontally, as demonstrated by the magnetic field lines in this region. The accompanying emissivity would be slightly stronger, resulting in the penumbra appearing brighter. As such, the emissivity in layered metallic hydrogen appears to be highly dependent on the orientation of the hexagonal hydrogen planes.

Likewise, it has been observed that sound waves travel faster within sunspots than within the photosphere [116, 117]. These findings are supportive of the idea that sunspots are denser and more metallic than the photosphere itself. The use of condensed matter brings with it both structure and function.

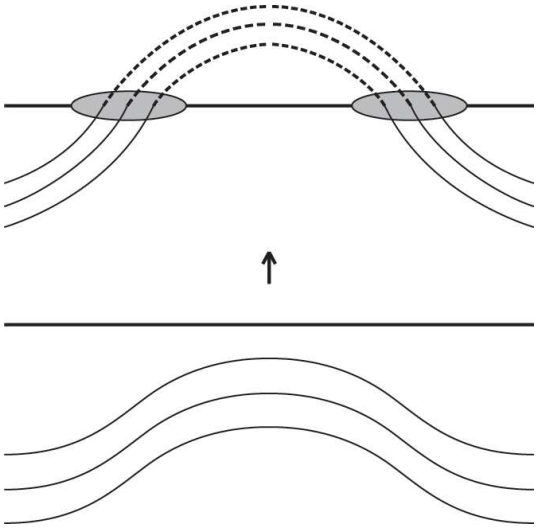

Fig. 8: Schematic representation of the appearance of a pair of sunspots on an active solar surface. The horizontal thick line illustrates the location of the photosphere, the thin lines the layers of metallic hydrogen, and the dashed lines the magnetic field. The two shaded circles outline the position of sunspots. In the lower portion of the figure, the layers of metallic hydrogen are below the level of the photosphere, but are being pushed up by intercalate elements which have entered the gas phase (see §5.1 in [48]). In the upper portion of the figure, the layers of metallic hydrogen have now broken through the photospheric level. The two sunspots are being linked solely by magnetic field lines, as the metallic hydrogen which once contained them has vaporized into the solar atmosphere. This figure is an adaptation based on Fig. 22 in [115]. Along with this legend, it previously appeared in [52].

2.3.4 Granular Emissivity #4

When observed at modest resolution, the surface of the Sun is covered with granules (see Fig. 9).* The appearance of these structures caused considerable controversy within astronomy in the mid-1800s [40], but they have been well described and illustrated [118–122] since the days of Father Secchi [1, p. 48–59, V. I]. Individual granules have limited lifetimes, can be arranged in mesogranules, supergranules, or giant cell [40, 118–122], and seem to represent a convective process.†

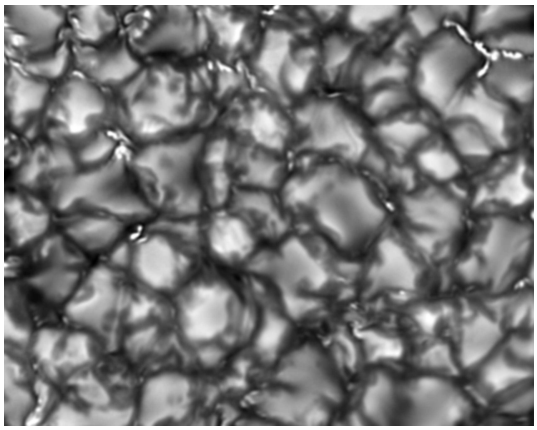

Fig. 9: High resolution image of solar granules acquired by Vasco Henriques on May 23, 2010 using the Swedish 1-m Solar Telescope (SST). Bright granules are surrounded by dark intergranular lanes which can contain magnetic bright points (see §2.3.5). This image has been described as follows, *"The SST is operated on the island of La Palma by the Institute for Solar Physics of the Royal Swedish Academy of Sciences in the Spanish Observatorio del Roque de los Muchachos of the Instituto de Astrofísica de Canarias — High resolution granulation — Observer: Vasco Henriques; Image processing: Vasco Henriques Date: 23 May 2010"*. http://www.solarphysics.kva.se (accessed online 9/15/2013).

Though granules are dynamic convective entities which are constantly forming and dying on the surface of the Sun, they have been found to observe the laws of Aboav-Weaire and of Lewis [123–125], along with the perimeter law, for space filling structures in two dimensions [126]. That granules can be viewed as crystals was first hypothesized by Chacornac in 1865 [127]. Clearly, the laws of space filling cannot be applied to gases which expand to fill the space of containers. They cannot, on their own, restrict the spatial extent which they occupy. The laws of space filling can solely be observed by materials which exist in the condensed state. Adherence to these laws by granules [126] constitutes important evidence that these structures are comprised of condensed matter.

Studies reveal that granules can contain *'dark dots'* at their center, linked to *'explosive'* structural decay. Rast [128] has stated that this decay *"can be better understood if granu-*

*The fourth Planckian proof [45] was initially part of the 14th line of evidence [45]. It has been presented, in greater detail, within [40] which contains an extensive list of references on the subject.

†This aspect of solar granules will be discussed in §5.1 as it is linked to activity on the solar surface. For the time being, the focus will remain on the structural and emissive aspects.

lation is viewed as downflow-dominated-surface-driven convection rather than as a collection of more deeply driven upflowing thermal plumes". These arguments depend on the presence of a true solar surface. Noever has linked the decay of granules associated with the appearance of *'dark dots'* to the perimeter law alone [126], once again implying that structure determines dynamic evolution.

Granules are characterized by important emissive characteristics. These structure tend to be brighter at their center and surrounded by dark intergranular lanes (see Fig. 9) whose existence has been recognized by the mid-1800s [40].

In order to account for the emissive properties of granules, the gaseous models maintain that these structures represent convective elements. Hot gases, rising from deep within the Sun, emerge near the center of these formations, while cooler material, held in the dark intergranular lanes, slowly migrates towards the solar interior. In this case, emissivity is linked to temperature changes alone, as dictated by Wien's law [17]. This hypothesis rests on the validity of Kirchhoff's law [15, 16, 20–24] and depends upon subtle changes in solar opacity [42] in adjacent regions of the solar surface. As seen in §2.1 and §2.2, these arguments are invalid.

Within the context of the LMH model [35, 39], granules are viewed as an integral portion of the true undulating surface of the Sun. Their complex radiative properties can be fully explained by considering directional spectral emissivity. As sub-components of the photosphere, the same mechanism invoked to understand limb darkening §2.3.2 can be used to explain granular emissivity.

The normal emissivity of these bubble-like structures remains somewhat elevated. As the viewing angle moves away from the normal,* emissivity progressively drops in accordance with the known behavior of non-metals (see Fig. 6). Intergranular lanes appear dark, not because they are cooler (an unlikely scenario in the same region of the Sun), but rather, because less photons are observed when the surface being visualized becomes increasingly coincident with the direction of emission. In a sense, with respect to thermal emission, each granule constitutes a mini-representation of the macroscopic limb darkening observed across the disk of the Sun (see §2.3.2), an idea first expressed by Very [101].

In the LMH model, granules therefore possess a Type I lattice [35, 39], which is somewhat less metallic than the Type-II lattice found in sunspots. This is revealed by the lack of strong magnetic fields associated with granules and by the slowly decaying center-to-limb variation in directional emissivity observed on the solar surface (see §2.3.2). In a manner analogous to what is observed in sunspots, the emissivity of layered metallic hydrogen would imply that the hexagonal hydrogen planes are oriented parallel to the solar surface at the center of a granules providing higher emissivity, or brighter appearance, in this instance. The orientation should become more vertical in the intergranular lanes, thereby accounting for their darker appearance. The LMH model [35, 39] dispenses with optical depth and variable temperature arguments. It elegantly accounts for solar emission using a single phenomenon (directional spectral emissivity in condensed matter) applicable across the full range of solar observations.

2.3.5 Facular Emissivity #5

In visible light, faculae are difficult to observe at the center of the solar disk, but often become quite apparent towards the limb.† Father Secchi noted the difficulty of observing faculae at the center of the disk [1, p. 49, V. I] and George Ellery Hale commented on the enhanced emissivity of faculae towards the limb, *"The bright faculae, which rise above the photosphere, are conspicuous when near the edge of the Sun, but practically invisible when they happen to lie near the center of the disk . . . "* [129, p. 85–86]. Solar faculae appear to float on the photosphere itself. The structures have long been associated with sunspots [130]. Wang et al. recently postulated that these objects could result from the conversion of sunspots, wherein the horizontal magnetic field contained within penumbrae makes a transition to a vertical field in faculae [131]. Faculae are known to possess strong magnetic fields [132–134].

The emissivity of faculae as they approach the solar limb [135] cannot be reasonably explained within the context of the gaseous models. The accepted scheme, Spruit's *'hot wall'* [136, 137] model is illustrated in Fig. 10. When the faculae are at the center of the disk, the observer is able to see deeper into the Wilson depression to the flux tube *'floor'* [137, p. 926]. This floor is thought to be at a lower temperature and, according to the laws of blackbody emission [15–20], appears relatively dark. As for the *'walls'* of the flux tube, they are said to sustain elevated temperatures and appear bright when compared to the deeper *'floor'*. As the flux tube moves towards the limb, the observer can no longer observe the *'floor'* and one of the *'hot walls'* becomes increasingly visible. With time, even that *'hot wall'* disappears. This agrees with observation: facular emissivity is initially indistinguishable from that of the photosphere at disk center. It then increases and becomes bright with respect to the rest of the solar surface, as theses objects move towards the limb. Finally, the emissivity decreases precipitously at the limb.

To help explain the emissivity of faculae, the gas models suggest macroscopic structures, *'cool floors'* and *'hot walls'*. Gases are incapable of generating such features. In faculae, flux tubes are said to be permitting heat from the solar interior to rise into the *'hot walls'*. Yet, to account for the darkness

*Normal viewing occurs when the line of sight is perpendicular to the surface.

†The fifth Planckian proof, as related to facular emissivity, was initially presented as the 15th line of evidence [45].

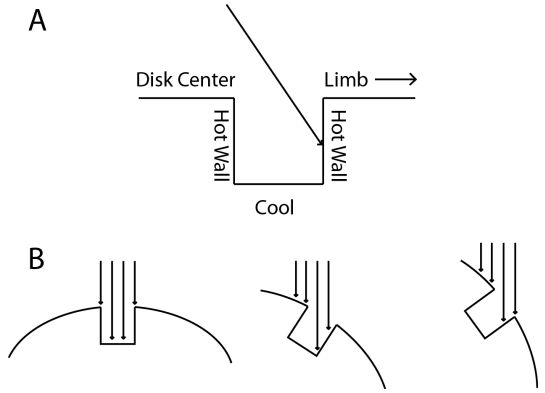

Fig. 10: Schematic representation of Spruit's *'hot wall'* model [136, 137]. A) Faculae are represented as depressions in the solar surface. Depending on the line of sight, the observer will sample either a *'cool floor'*, or a *'hot wall'*. B) When sampling at the center of the solar disk, he/she will only be able to visualize a *'cool floor'* whose temperature approaches that of the granules on the surface. Under the circumstances, the faculae are not visible. However, as these objects move towards the limb, the line of sight will initially sample more of the *'hot wall'* and the faculae appear brighter. When the edge of the Sun is approached, the hot walls can no longer be readily sampled and the emissivity of the faculae are perceived to drop rapidly.

within sunspots, the models had required that field lines inhibited the upward flow of hot gases beneath the umbra (see §2.2.3).

It is immediately apparent that the emissive behavior just described within faculae exactly parallels the known radiative properties of metals, as previously illustrated in Fig. 6. Faculae possess strong magnetic fields [132–134]. In combination with their directional emissivity, this all but confirms that they are metallic in nature.

In addition to faculae, an extension of Spruit's hot wall model has been invoked to explain the presence of magnetic bright points found within the dark intergranular lanes of the granules [138]. As the name implies, magnetic bright points are also believed to possess strong magnetic fields [12, 138, 139]. Moreover, they display powerful center-to-limb variations in their emissivity [138], being most visible at the center of the solar disk within the dark intergranular lanes. In the case of magnetic bright points, it is the *'floor'* which is viewed as bright, as light is said to originate from *"deeper photospheric layers that are usually hotter"* [138].*

The problem rests in the realization that magnetic bright points are located within the dark intergranular lanes. As a result, in order to explain the presence of locally strong magnetic fields within these objects, it is hypothesized that an *"efficient turbulent dynamo transforms into magnetic fields part of the kinetic energy of the granular convection"* [138]. This serves to emphasize the problems faced by the gas models.

Within the context of the LMH model [35, 36, 39], the presence of faculae and magnetic bright points on the solar surface are elegantly explained by invoking lattice structure. Since faculae are associated with sunspots [130] and even thought to be ejected from these structures [131], it is reasonable to propose that they can be metallic in nature (see Fig. 6), that their structural lattice mimics the type II lattice found in sunspots, and that they have not yet relaxed back to the Type-I lattice found in granules. In this case, the brightness of faculae implies that their hexagonal hydrogen planes lie parallel to the solar surface. This should account for both emissivity and the presence of associated magnetic fields in these structures.

In the end, the simplest explanation for the origin for magnetic bright points may be that they are nothing more than facular elements. Rising from internal solar regions, they have not fully relaxed from a Type II to a Type I lattice, but have been transported through granular flow to deeper intergranular lanes. Their center-to-limb emissivity variations may well rest in the realization that they are hidden from view by the granules themselves as the limb is approached. Hence, their numbers appear to fall towards the edge of the solar disk [138].

2.3.6 Chromospheric Emissivity #6

While hydrogen-α emissions are responsible for the red glow of the chromosphere visible during an eclipse, this region of the Sun also emits a weak continuous spectrum [56] which has drawn the attention of solar observers for more than 100 years [140–147].† Relative to this emission, Donald Menzel noted, *"...we assumed that the distribution in the continuous chromospheric spectrum is the same as that of a black body at 5700°, and that the continuous spectrum from the extreme edge is that of a black body at 4700°. There is evidence in favor of a lower temperature at the extreme limb in the observations by Abbot, Fowle, and Aldrich of the darkening towards the limb of the Sun"* [142].

The gaseous models infer that the chromosphere has an average density of $\sim 10^{-12}$ g/cm^3 [115, p. 32].‡ Despite a 10^5 drop in density with respect to the photosphere, these treatments continue to advance that the continuous emission in the chromosphere is being produced by neutral H, H$^-$, Rayleigh scattering, and electron scattering (see [145, 146] and [150, p. 151–157]). But, none of these processes can be found in graphite (see §2.1 and §2.2).

*These layers were not hotter in Spruit's model [136, 137].

†The sixth Plankian proof [45] was initially presented as the 26th line of evidence [56].

‡In these models, the photosphere is assumed to have a density of $\sim 10^{-7}$ g/cm^3, while the outer chromosphere has a density of $\sim 10^{-15}$ g/cm^3 [148]. This constitutes an 8 order of magnitude decrease in just a few thousand kilometers. As a point of reference, the density of the Earth's atmosphere at sea level is $\sim 1.2 \times 10^{-3}$ g/cm^3 [149] or ~10,000 greater than calculated photospheric densities for the gas models.

Alternatively, within the context of the LMH model, the chromospheric continuous emission provides evidence that condensed matter exists in this region of the solar atmosphere [56]. This is in keeping with the understanding that continuous spectra, which can be described using blackbody behavior, must be produced by condensed matter [21–24]. In this regard, the chromosphere may be viewed as a region of hydrogen condensation and recapture within the Sun. Though generating condensed matter, the chromosphere is not comprised of metallic hydrogen.*

2.3.7 K-Coronal Emissivity #7

The white light emitted by the K-corona is readily visualized during solar eclipses.† Observing from Iowa in 1869, William Harkness *"obtained a coronal spectrum that was continuous except for a single bright green line, later known as coronal line K1474"* on the Kirchhoff scale [151, p. 199]. Eventually, it became clear that the continuous spectrum of the K-corona was essentially identical to photospheric emission [152–156], with the important distinction that the former was devoid of Fraunhofer lines. In addition, the spectrum of the K-corona appeared to redden slightly with increasing distance from the solar surface, *"microphotograms for solar distances varying from R=1.2s to R=2.6s show that the coronal radiation reddens slightly as the distance from the Sun is increased"* [156]. The reddening of the K-coronal emission suggested that the corona was cooling with increased distance from the solar surface.‡

Within the context of the gas models, the corona is extremely hot and thus, cannot be self-luminous in the visible spectrum. Rather, these models maintain that coronal white light must represent photospheric radiation. But as the thermal spectrum from the photosphere is punctuated with Fraunhofer absorption lines (see §3.7), some mechanism must be devised to explain their absence in coronal light. As such, proponents of the gaseous models have proposed that coronal light is being scattered by highly relativistic electrons [115, 148, 157, 158]. The Fraunhofer absorption lines are hypothesized to become highly broadened and unobservable. Relativistic electrons require temperatures in the millions of degrees. These temperatures are inferred from the line emissions of highly ionized ions in this region of the Sun (see §3.8). Unfortunately, such a scheme fails to account for the reddening of the coronal spectrum [156].

In contrast, the LMH model [35, 39] states that the solar corona contains photospheric-like condensed matter (Type I) and is, accordingly, *self-luminous* [57]. It is well-known that the Sun expels material into its corona in the form of flares and coronal mass ejections. It is reasonable to conclude that this material continues to emit (see §2.3.8) and may eventually disperse into finely distributed condensed matter in this region of the Sun. The reddening of the coronal spectrum implies that the apparent temperatures of the corona are no greater than those within the photosphere.§ The apparent temperature slowly decreases, as expected, with increased distance from the solar surface. The production of highly ionized ions in the corona reflects condensed matter in the outer solar atmosphere (see §2.3.8, §3.8, and §5.5). As for the Fraunhofer lines, they do not appear on the spectrum of the K-corona owing to insufficient concentrations of absorbing species exist in this region of the Sun. There is no need to invoke scattering by relativistic electrons.

2.3.8 Coronal Structure Emissivity #8

The corona of the active Sun is filled with structures easily observed using white-light coronographs [154, 155].¶ Flares [159–162], prominences and coronal mass ejections [163–171], streamers [172–174], plumes [175], and loops [176–178], can all be visualized in white light.

The mechanism for generating white-light in this wide array of structures remains elusive for the gaseous models, in part because the densities, in which they are hypothesize to exist, are lower than $\sim 10^{-15}$ g/cm^3 [148]. Moreover, the release of white-light by these structures tends to be explosive in nature, particularly when flares are involved [179–186]. These phenomena cannot be adequately explained by relying on gradual changes in opacity [42] or the action of rela-

*Metallic hydrogen requires extreme pressures for formation [39, 92] which can only exist within the solar body. As a result, though condensation is occurring within the chromosphere and corona, the resulting products are not metallic. Rather, it is likely that chromospheric material is comprised of dense hydrogen wherein molecular interactions between hydrogen atoms still persists [92]. Conversely, condensed matter which has been ejected from the solar body can be metallic in character and has been proposed to become distributed throughout the corona [60]. The solar atmosphere can simultaneously support the existence of two forms of hydrogen: chromospheric non-metallic material, like as coronal rain or spicules (see §5.4, §5.6 and [53, 59]) and coronal material which resembles photospheric Type-I metallic hydrogen (see §2.3.7 and §2.3.8) and [57, 58, 60]) and which can be found in the corona and its associated structures (see §3.8, §4.6, §5.5, §5.7 and §6.6 for complimentary evidence).

†The seventh Plankian proof [45] was initially presented as the 27th line of evidence [57, 60].

‡Yet, the *"single bright green line"* which had been observed by Harkness would eventually be identified as originating from highly ionized iron (i.e. FeXIV). Within the gaseous context, the only means of generating these ions would involve the presence of extreme temperatures in the corona. Conversely, the ions could be produced if condensed matter can be postulated to exist in this region of the Sun. The origin of highly ionized ions in the corona constitutes one of the most elegant lines of evidence for the presence of condensed matter in this region of the Sun, supporting the idea that the corona is, in fact, cool (see [60] and §3.8 for a complete discussion).

§The author has stated that the true energy content of the photosphere would correspond to real temperatures in the millions of degrees. The vast majority of this energy is trapped within the translational degrees of freedom associated with the differential convection currents. The conduction bands responsible for the solar magnetic fields likewise harness some of the solar surface energy. The apparent temperature of ~6,000K corresponds to the energy contained within the photospheric vibrational degrees of freedom [41].

¶The eighth Plankian proof [45] was initially presented as the 28th line of evidence [58].

tivistic electrons to scatter photospheric light [160, 161, 164, 187, 188]. Currently, many of these structures are believed to derive their energy from coronal magnetic sources overlying active regions [12]. That is a result having no other means of accounting for this extensive and abrupt release of energy in the gaseous Sun [179].

Within the context of the LMH model [35, 39], the white-light emitted by coronal structures is associated with their condensed nature. Since many of these formations originate from eruptions taking place at the level of the photosphere, such a postulate appears reasonable. As a result, coronal structures should be regarded as *self-luminous*. The explosive increase in white-light is related to powerful lattice vibrations associated with their formation [21]. Long ago, Zöllner [189] had insisted that flares involved the release of pressurized material from within the Sun [3]. These mechanisms remain the most likely, as they properly transfer energy out of the solar body, not back to the surface from the corona (see §5.1).

3 Spectroscopic Lines of Evidence

Though Gustav Kirchhoff erred [21–24] relative to his law of thermal emission [15, 16], his contributions to solar science remain unchallenged. Not only was he amongst the first to properly recognize that the Sun existed in liquid state [2, 26], but as the father of spectral analysis, along with Robert Bunsen, he gave birth to the entire spectroscopic branch of solar science [190, 191]. Using spectroscopic methods, Kirchhoff successfully identified the lines from sodium on the Sun and this led to an avalanche of related discoveries, spanning more than a century [190, 191]. Indeed, all of the thermal proofs discussed in §2, are the result of spectroscopic analysis, centered on the blackbody spectrum observable in visible and infrared light. It is fitting that the next series of proofs are spectroscopic, this time centering on line emission of individual atoms or ions. These eight lines of evidence highlight anew the power of Kirchhoff's spectroscopic approaches.

3.1 UV/X-ray Line Intensity #9

The Sun is difficult to study in the ultraviolet (UV) and X-ray bands due to the absorption of this light by the Earth's atmosphere.* As a consequence, instruments like the AIA aboard NASA's Solar Dynamic Observatory (see Fig. 11) are being used for these observations [192, p. ix]. When the Sun is observed at these frequencies, striking evidence is produced on the existence of a real solar surface. Harold Zirin describes the findings as follows, *"The case in the UV is different, because the spectrum lines are optically thin. Therefore one would expect limb brightening even in the absence of temperature increase, simply due to the secant increase of path length. Although the intensity doubles at the limb, where we see the back side, the limb brightening inside the limb is minimal . . . Similarly, X-ray images show limb brightening simply due to increased path length."* [193]. Fig. 11 presents this phenomenon in X-Ray at 94Å, for a somewhat active Sun.†

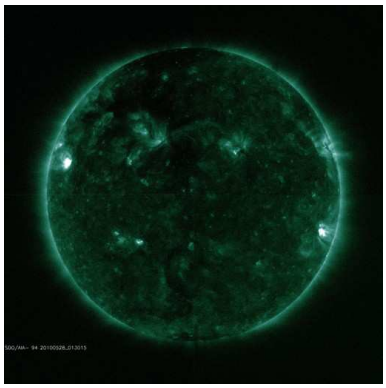

Fig. 11: AIA X-Ray image of an active Sun obtained on 5/28/2010 at 94Å displaying limb brightening and surface activity. This image (20100528_013015_512_0094.jpg) has been provided Courtesy of NASA/SDO and the AIA, EVE, and HMI science teams using data retrieval (http://sdo.gsfc.nasa.gov/data/aiahmi).

When the observer is directly examining the center of the opaque solar disk, weak spectral lines are obtained at these frequencies. The lines brighten slightly as observation moves towards the limb, owing to a slightly larger fraction of the solar atmosphere being sampled (line of sight 2 versus 1 in Fig. 12). However, immediately upon crossing the solar limb, a pronounced increase in spectroscopic intensity can be recorded. In fact, it approximately doubles, because a nearly two-fold greater line of sight is being viewed in the solar atmosphere. This can be understood if one would compare a line of sight very near line 3 in Fig. 12 (but still striking the solar disk) with line 3 itself.

In this manner, UV and X-ray line intensities can provide strong evidence that the Sun possesses an opaque surface at these frequencies which is independent of viewing angle. Limb darkening is not observed, as was manifested in the visible spectrum (see §2.3.2), in that condensed matter is not being sampled. Rather, the behavior reflects that gases are being monitored above a distinct surface through which UV and X-ray photons cannot penetrate.‡

3.2 Gamma-Ray Emission #10

Occasionally, powerful gamma-ray flares are visible on the surface of the Sun and Rieger [194] has provided evidence that those with emissions >10 MeV are primarily visualized

*This proof was first presented as the 25th line of evidence [55].

†A 171Å UV image from the quite Sun has been published [192, p. 38]. The Solar Dynamic Observatory website can be accessed for images at other frequencies in the ultra-violet (http://sdo.gsfc.nasa.gov/data/aiahmi).

‡Note that these findings further bring into question the optical depth arguments that had been brought forth to explain limb darkening within the gaseous models in §2.3.3. Should the Sun truly possess a vacuum-like photospheric density of only 10^{-7} g/cm^3 [148], then the limb should not act as such a dramatic boundary relative to the intensity of UV and X-ray emissions.

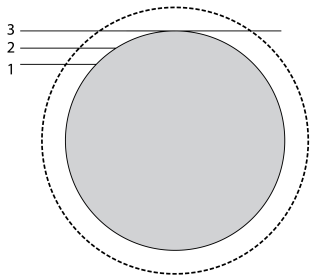

Fig. 12: Schematic representation of path lengths present when the outer atmosphere (area outlined by dashes) of the Sun (body in gray) is viewed from the Earth. Paths 1 and 2 terminate on the solar surface. Just beyond the limb, path 3 samples the front and back side of the solar atmosphere, resulting in a two fold increase in line intensity. This figure is an adaptation based on Fig. 2.4 in [192] and, along with this legend, was previously published [55].

near the solar limb (see Fig. 13).* Speaking of Rieger's findings, Ramaty and Simnett noted that "*Gamma-ray emitting flares are observed from sites located predominantly near the limb of the Sun ... This effect was observed for flares detected at energies >0.3 MeV, but it is at energies >10 MeV that the effect is particularly pronounced ... Since in both of these cases the bulk of the emission is bremsstrahlung from primary electrons, these results imply that the radiating electrons* (are) *strongly anisotropic, with more emission in the directions tangential to the photosphere than in directions away from the Sun*" [195, p. 237].

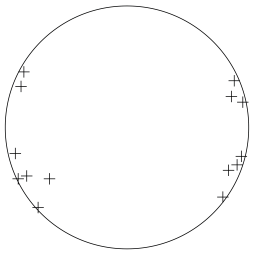

Fig. 13: Schematic representation of approximate flare positions with >10 MeV of energy on the solar disk displaying their predominance near the limb. This figure is meant only for illustrative purposes and is an adaptation based on Fig. 9 in [194] which should be examined for exact flare locations. This figure was previously published in [49].

The production of anisotropic emission would typically imply that structural constraints are involved in flare production. Since the gaseous Sun cannot sustain structure, another means must be used to generate this anisotropy. Based on theoretical arguments, Ramaty and Simnett consequently advance that: "*... the anisotropy could result from the mirroring of the charged particles in the convergent chromospheric magnetic fields*" [195, p. 237]. The anisotropy of gamma-ray emission from high energy solar flares is thought to be generated by electron transport in the coronal region and magnetic mirroring of converging magnetic flux tubes beneath the transition region [195]. The energy required for flare generation could thereby be channeled down towards the solar surface from the corona itself. Conveniently, the chromosphere instantly behaves as an *'electron mirror'*. Devoid of a real surface, another mechanism was created to *act* as a surface.

The inability to generate flare anisotropy using the most obvious means — the presence of a true photospheric surface — has resulted in a convoluted viewpoint. Rather than obtain the energy to drive the flare from within the solar body, the gaseous models must extract it from the solar atmosphere and channel it down towards the surface using an unlikely mechanism. It remains simpler to postulate that the anisotropy observed in high energy solar flares is a manifestation that the Sun has a true surface. The energy involved in flare generation can thereby arise from the solar interior, as postulated long ago by Zöllner [189]. In this respect, the LMH model [35, 39] retains distinct advantages when compared to the gaseous models of the Sun.

3.3 Lithium Abundances #11

Kirchhoff's spectroscopic approaches [190, 191] have enabled astronomers to estimate the concentrations of many elements in the solar atmosphere.† Application of these methods have led to the realization that lithium was approximately 140-fold less abundant in the solar atmosphere than in meteors [196, 197].

In order to explain this discrepancy, proponents of the gaseous stars have advanced that lithium must be transported deep within the interior of the Sun where temperatures $>2.6\times10^6$ K are sufficient to destroy the element by converting it into helium [$^7\mathrm{Li}\,(p,\alpha)^4\mathrm{He}$] [198]. To help achieve this goal, lithium must be constantly mixed [198–200] into the solar interior, a process recently believed to be facilitated by orbiting planets [201, 202]. Though these ideas have been refuted [203], they highlight the difficulty presented by lithium abundances in the gaseous models.

As for the condensed model of the Sun [35, 39], it benefits from a proposal [54], brought forth by Eva Zurek, Neil Ashcroft, and others [204], that lithium can act to stabilize metallic hydrogen [88, 92]. Hence, lithium levels could appear to be decreased on the solar surface, as a metallic hydrogen Sun retains the element in its interior. At the same time, lithium might be coordinated by metallic hydrogen in the corona, therefore becoming sequestered and unavailable for emission as an isolated atom.

*This proof was first presented as the eighteenth line of evidence [49].

†This proof was initially discussed in [54]. See [47], for a detailed discussion of how elemental abundances have been estimated.

In this manner, lithium might be unlike the other elements, as these, including helium, are likely to be expelled from the solar interior (see §5.1) as a result of exfoliative forces [48]. Lithium appears to have a low abundance, but, in reality, it is not being destroyed. This would better reconcile the abundances of lithium observed in the solar atmosphere with that present in extrasolar objects. Clearly, if lithium is being destroyed within the stars, it becomes difficult to explain its abundance in meteors. This problem does not arise when abundances are explained using a LMH model, as metallic hydrogen can sequester lithium into its lattice.

3.4 Hydrogen Emission #12

The *'flash spectrum'* associated with solar eclipses characterizes the chromosphere.* The strongest features within this spectrum correspond to line emissions originating from excited hydrogen atoms. As far back as 1931, the outstanding chromospheric observer, Donald H. Menzel, listed more than twenty-three hydrogen emission lines originating from this region of the Sun (see Table 3 in [205, p. 28]). It is the cause of these emissions which must now be elucidated. The most likely scenario takes advantage of the condensation appearing to occur in the chromospheric layer (see §5.4 §5.6 and [56, 59]).

By modern standards, the nature of the chromosphere remains a mystery, as Harold Zirin reminds us, *"The chromosphere is the least-well understood layer of the Sun's atmosphere...Part of the problem is that it is so dynamic and transient. At this height an ill-defined magnetic field dominates the gas and determines the structure. Since we do not know the physical mechanisms, it is impossible to produce a realistic model. Since most of the models ignored much of the data, they generally contradict the observational data. Typical models ignore other constraints and just match only the XUV data; this is not enough for a unique solution. It reminds one of the discovery of the sunspot cycle. While most of the great 18th century astronomers agreed that the sunspot occurrence was random, only Schwabe, an amateur, took the trouble to track the number of sunspots, thereby discovering the 11-year cycle"* [193]. But if mystery remains, it is resultant of the denial that condensed matter exists in this layer of the Sun.

The chromosphere is characterized by numerous structural features, the most important of which are spicules (see Fig. 14) [59, 150]. Even in the mid-1800s, Secchi would provide outstanding illustrations of these objects (see Plate A in [1, V. II]). He would discuss their great variability in both size and orientation, *"In general, the chromosphere is poorly terminated and its external surface is garnished with fringes ... It is almost always covered with little nets terminated in a point and entirely similar to hair ... it often happens, especially in the region of sunspots, that the chromosphere presents an aspect of a very active network whose surface, unequal and rough, seems composed of brilliant clouds analogous to our cumulus; the disposition of which resembles the beads of our rosary; a few of which dilate in order to form little diffuse elevations on the sides"* [1, p. 31–36, V. II].

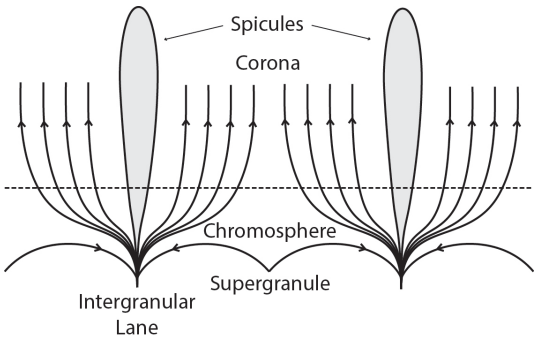

Fig. 14: Schematic representation of spicules overlying the intergranular lane on the outer boundary of a supergranule and surrounded by magnetic field lines emanating from the solar surface. While simplistic, this illustration conveys the basic structural elements needed for discussion. This figure was previously published in [59] and is an adaptation based on Fig. IV-13 in [206, p. 162].

At first glance, spicules are thought to have a magnetic origin, as these fields seem to flood the chromosphere [148, 150, 206–215]. In reality, matter within the chromosphere seems to form and dissipate quickly and over large spatial extent, with spicules reaching well into the corona [148, 150, 206–215]. The random orientation which spicules display, as noted long ago by Secchi [1, p. 31–36, V. II], along with their velocity profiles (see §5.6), should have dispelled the belief that these structures are magnetic in origin. Rather, they appear to be products of condensation (§5.6).†

If spicules and chromospheric matter are genuinely the product of condensation reactions, then their mechanism of formation might shed great light into the emissive nature of this solar layer.

3.4.1 The Liquid Metallic Hydrogen Solar Model

The search for answers begins by considering condensation processes known to occur on Earth [59].

In this respect, while studying the agglomeration of silver clusters, Gerhart Ertl's (Nobel Prize, Chemistry, 2007) laboratory noted that *"Exothermic chemical reactions may be accompanied by chemiluminescence. In these reactions, the released energy is not adiabatically damped into the heat bath of the surrounding medium but rather is stored in an ex-*

*This proof was first presented as the seventeenth line of evidence [47, 59].

†While non-magnetic, spicules might nonetheless be confined by magnetic fields present in the charged plasmas or coronal metallic hydrogen that surrounds them, much as illustrated in Fig. 14.

cited state of the product; decay from this excited state to the ground state is associated with light emission" [216].

The reactions of interest are seldom studied. Those which must arouse attention involve the condensation of two silver fragments and the formation of an activated cluster species: $Ag_n + Ag_m \rightarrow Ag^*_{m+n}$ [216]. With respect to the chromosphere, the important features of these reactions involve the realization that condensation processes are exothermic.

When silver clusters condense, energy must be dissipated through light emission. This constitutes a vital clue in explaining why the chromosphere is rich in hydrogen emission lines [59, 205]. Once an activated cluster is formed, it can relax by ejecting an excited atom: $Ag^*_{m+n} \rightarrow Ag_{m+n-1} + Ag^*$. The reactions are completed when the ejected excited species emits light to reenter the ground state: $Ag^* \rightarrow Ag + h\nu$.

Taking guidance from the work in metal clusters [216], hydrogen emission lines in the chromosphere might be seen as produced through the condensation of hydrogen fragments, $H_n + H_m \rightarrow H^*_{m+n}$. The resultant condensation product could then relax through the ejection of an excited hydrogen atom, $H^*_{m+n} \rightarrow H_{m+n-1} + H^*$, which finally returns to a lower energy state with light emission, $H^* \rightarrow H + h\nu$. This could give rise to all the Lyman lines ($N_2 > 1 \rightarrow N_1 = 1$). If one postulates that the excited hydrogen atom can hold its electron in any excited orbital $N_2 > 2$, H^{**}, then the remaining complement of hydrogen emission lines could be produced $H^{**} \rightarrow H^* + h\nu$ (Balmer $N_2 > 2 \rightarrow N_1 = 2$, Paschen series $N_2 > 3 \rightarrow N_1 = 3$, and Brackett series $N_2 > 4 \rightarrow N_1 = 4$).

But since the chromosphere is known to possess spicules and mottles [148, 150, 206–215], it is more likely that hydrogen is condensing, not onto a small cluster, but rather, onto very large condensed hydrogen structures, CHS [59].* The most logical depositing species in these reactions would be molecular hydrogen, as it has been directly observed in sunspots [217, 218], on the limb [219], and in flares [218]. Importantly, the emission from molecular hydrogen is particularly strong in chromospheric plages [220], providing further evidence that the species might be the most appropriate to consider.

As a result, it is reasonable to postulate that molecular hydrogen could directly interact with large condensed hydrogen structures, CHS, in the chromosphere [59]. The reaction involved would be as follows: $CHS + H_2 \rightarrow CHS\text{–}H^*_2$. This would lead to the addition of one hydrogen at a time to large condensed structures and subsequent line emission from the ejected excited species, $H^* \rightarrow H + h\nu$. Numerous reactions could simultaneously occur, giving rise to the rapid growth of chromospheric structures, accompanied with significant light emission in all spectral series (i.e. Lyman, Balmer, Paschen, and Brackett).

*Chromospheric condensed hydrogen structures, CHS, are likely to be composed of extremely dense condensed matter wherein molecular hydrogen interactions linger [92].

3.4.2 The Gaseous Solar Models

The situation being promoted in §3.4.1, concerning hydrogen line emission in the chromosphere, is completely unlike that currently postulated to exist within the gaseous Sun [59]. In the gas models, line emission relies on the accidental excitation of hydrogen through bombardment with either photons or electrons [206, p. 2]. The process has no purpose or reason. Atoms are randomly excited, and then, they randomly emit.

Przybilla and Butler have studied the production of hydrogen emission lines and the associated lineshapes in the gaseous models. They reached the conclusion that some of the hydrogen emission lines *"collisionally couple tightly to the continuum"* [221]. Their key source of opacity rests with the H^- ion, which has previously been demonstrated to be incapable of providing the desired continuous emission [42]. Of course, it is impossible to *"collisionally couple tightly to the continuum"* [221] in the gaseous models, as the continuum originates solely from opacity changes produced by an array of processes [42]. In the chromosphere, where average densities are postulated to be extremely low ($\sim 10^{-15}$ g/cm^3 [148]), continuous emission is thought to be produced by neutral H, H^-, Rayleigh scattering, and electron scattering (see [145, 146] and [150, p. 151–157]). Clearly, it is not possible to tightly couple to all of these mechanisms at once.

Przybilla's and Butler's computations [221] involve consideration of line blocking mechanisms and associated opacity distribution functions [222]. Stark line broadening mechanisms must additionally be invoked [223].

Beyond the inability of gases to account for the continuous spectrum and the shortcomings of solar opacity calculations [42], the central problem faced in trying to explain hydrogen emission and the associated line shapes rests in the Stark mechanisms themselves. Stark line broadening relies upon the generation of local electric fields near the emitting hydrogen atom. These fields are believed to be produced by ions or electrons which come into short term contact with the emitting species [223]. On the surface at least, the approach seems reasonable, but in the end, it relies on far too many parameters to be useful in understanding the Sun.

In the laboratory, Stark broadening studies usually center upon *extremely dense plasmas*, with electron numbers approaching 10^{17} cm^{-3} [224]. Stehlé, one of the world's preeminent scientists relative to Stark linewidth calculations [223, 225, 226], has analyzed lineshapes to infer electron numbers ranging from 10^{10} to 10^{17} cm^{-3} [227].† She initially assumes that plasmas existing within the chromosphere (T=10,000 K) have electron numbers in the 10^{13} cm^{-3} range [223]. Other

†While the vast majority of plasma studies report electron densities in the 10^{17} cm^{-3} range, the He I studies range from 10^{15} cm^{-3} to 10^{17} cm^{-3} [224]. The lowest electron numbers, 10^{15} cm^{-3}, are produced using arc discharge low density plasma settings. However, these could have little relevance in the Sun, as arc experiments rely on the capacitive discharge of large voltages. They do not depend on fluctuating electromagnetic fields [228].

sources call for much lower values. For instance, electron numbers of $\sim 10^{16}$ m^{-3} (or $\sim 10^{10}$ cm^{-3}) are obtained from radio measurements by Cairns et al. [229] and of no more than $\sim 10^{15}$ m^{-3} (or $\sim 10^{9}$ cm^{-3}) are illustrated in Dwivedi Fig. 3 [157, p. 285]. Stark experiments on Earth typically utilize electron numbers which are approximately 1–100 million times greater than anything thought to exist in the chromosphere.

A minor objection to the use of Stark broadening to explain the width of the hydrogen lines in the gaseous models rests on the fact that the appropriate experiments on hydrogen plasma do not exist. The plasma form of hydrogen (H II) is made of protons in a sea of electrons. It lacks the valence electron required for line emission. The closest analogue to excited hydrogen in the Sun would be ionized helium in the laboratory [224], although ionized Argon has been used for the Hβ profile [227].*

However, the most serious problem rests in the realization that these methods are fundamentally based on the presence of electric or electromagnetic fields in the laboratory. For instance, the inductively produced plasmas analyzed by Stehlé [227] utilize discharges on the order of 5.8 kV [227]. Inductively produced plasmas involve directionally-oscillating electromagnetic fields. Spark or arc experiments utilize static electric fields to induce capacitive discharges across charged plates. In every case, the applied electric field has *a distinct orientation*. Such conditions are difficult to visualize in a gaseous Sun, particularly within the spicules (see §3.4 and §5.7), given their arbitrary orientations. Random field orientations are incapable of line broadening, as well understood in liquid state nuclear magnetic resonance.

Stark broadening requires constraints on the electric field. In the gaseous models, these must take the form of a charged particle which approaches, precisely at the correct moment, an emitting species. The use of such mechanisms to account for chromospheric line profiles is far from justified. But, as the gaseous models cannot propose another explanation, everything must rest on Stark mechanisms, however unlikely these are to be valid in this setting.

In the end, it is not reasonable that matter existing at the concentration of an incredible vacuum ($\sim 10^{-15}$ g/cm^3 [148]) could be Stark broadened, given the extremely low electron numbers associated with the chromosphere [157, 229]. Computations have merely extended our *'observational range'* to electron numbers never sampled in the laboratory. According to the gas models, the chromosphere is a region of extremely low density, but high density plasmas must be studied to enable Stark analysis. Then, while the results of Stark broadening calculations appear rigorous on the surface, they contain experimental shortcomings. Spatially aligned electric fields cannot exist throughout the spicular region of a fully gaseous solar atmosphere, lone electrons are unlikely to produce the desired electric fields, and atoms such as argon have little relevance to hydrogen. In any case, given enough computational flexibility, any lineshape can be obtained, but opacity considerations remain [42].

3.4.3 Summary

As just mentioned in §3.4.2, Stark experiments involve electron densities far in excess of anything applicable to the solar chromosphere. Using the same reasoning, it could be argued that metallic hydrogen has not been created on Earth [39, 92]. The criticism would be justified, but this may be simply a matter of time. Astrophysics has already adopted these materials in other settings [93–96] and experimentalists are getting ever closer to synthesizing metallic hydrogen [39, 92]. The Sun itself appears to be making an excellent case that it is comprised of condensed matter.

Unlike the situation in the gaseous solar models, where hydrogen emission becomes the illogical result of random reactions, within the context of the liquid hydrogen model, it can be viewed as the byproduct of systematic and organized processes (see §3.4.1). An underlying cause is associated with line emission, dissipation of the energy liberated during condensation reactions. The driving force is the recapture of hydrogen through condensation, leading ultimately to its re-entry into the solar interior. This tremendous advantage cannot be claimed by the gaseous models.

Pressure (or collisional) broadening can be viewed as the most common mechanism to explain line broadening in spectroscopy. This mechanism can be invoked in the condensed model, because the atmosphere therein is not devoid of matter (see §2.3.6, §5.4, §5.5, §5.6, §6.6 and [56, 58, 59]).

It is possible that line broadening is occurring due to direct interaction between the emitting species and condensed hydrogen structures in the chromosphere. In this case, emission would be occurring simultaneously with the ejection of hydrogen. Under the circumstances, hydrogen line shapes may be providing important clues with respect to the interaction between molecular hydrogen and larger condensed structures in the chromosphere. If Stark broadening mechanisms play any role in the Sun, it will only be in the context of condensed matter generating the associated electric field.

3.5 Elemental Emission #13

Beyond hydrogen, the solar chromosphere is the site of emission for many other species, particularly the metals of the main group and transition elements.† For gaseous models, these emissions continue to be viewed as the product of random events (see §3.4.2). However, for the LMH model, con-

*The use of argon to represent hydrogen immediately suggests that these methods are not relevant to the Sun. Unlike hydrogen, argon has valence shells containing up to 18 electrons. This many electrons, when either ionized or polarized, presents an analogue with little or no resemblance to hydrogen and its lone electron.

†This proof was first presented as the thirtieth line of evidence [59].

densation remains the focus (§3.4.1), but this time with the assistance of the hydrides.

The solar disk and the sunspots are rich in hydrides including CaH, MgH, CH, OH, H_2O, NH, SH, SiH, AlH, CoH, CuH, and NiH [230, 231]. CaH and MgH have been known to exist in the Sun for more than 100 years [232]. Hydrogen appears to have a great disposition to form hydrides and this is important for understanding the role which they play in the chromosphere.

At the same time, the emission lines from CaII and MgII are particularly strong in the chromosphere [206, p. 361-369]. These represent emissions from the Ca^+ and Mg^+ ions. Yet, the inert gas configurations for these atoms would lead one to believe that the Ca^{+2} (CaIII) and Mg^{+2} (MgIII) lines should have been most intense in the chromosphere. As such, why is the Sun amplifying the CaII and MgII lines? Surely, this cannot be a random phenomenon (§3.4.2),* as these should have led to the buildup of the most stable electronic configuration.

The answer may well lie in reconsidering the condensation reactions presented in §3.4.1, but this time substituting CaH for molecular hydrogen. It should be possible for CaH and a condensed hydrogen structure, CHS, to interact, thereby forming an activated complex, CHS + CaH → CHS-HCa*. This complex could then emit a CaII ion in activated state, Ca^{+*}, and capture the hydrogen atom: CHS–HCa* → CHS–H + Ca^{+*}. Finally, the emission lines from CaII would be produced, as Ca^{+*} (CaII*) returns to the ground state: $Ca^{+*} \rightarrow Ca^+ + h\nu$. As was the case when discussing the condensation of molecular hydrogen (§3.4.1), if one permits the electrons within the excited state of CaII to initially occupy any electronic orbital, CaII**, then all possible emission lines from CaII could be produced: $Ca^{+**} \rightarrow Ca^{+*} + h\nu$. A similar scheme could be proposed for MgH and the other metal hydrides, depending on their relative affinity for CHS.

There is an important distinction between this scenario and that observed with molecular hydrogen (§3.4.1). When metal hydrides are utilized in this scheme, the condensation reactions are delivering both a proton and *two* electrons to the condensed hydrogen structure. The reactions involving molecular hydrogen delivered a single electron. This interesting difference can help to explain the varying vertical extent of the chromosphere when viewed in Hα, CaII, or HeII (see §3.6 and §4.7).

When sampling the solar atmosphere, electron densities appear to rise substantially as one approaches the photosphere (see [229] and [157, p. 285]). Hence, the lower chromosphere is somewhat electron rich with respect to the upper regions of this layer. Thus, in the lower chromosphere, condensation reactions involving the ejection of atomic hydrogen and neutral atoms can abound. As the altitude increases, a greater affinity for electrons arises and condensation can now be facilitated by species like as the metal hydrides, which can deliver two electrons per hydrogen atom.† This explains why CaII lines in the chromosphere can be observed to rise to great heights [193].

At the same time, lines from neutral metals, M, are more prevalent in the lower chromosphere [193]. Since this area is electron rich, a two electron delivery system is unnecessary and reactions of the following form can readily occur: 1) MH + CHS → CHS-HM*, 2) CHS–HM* → CHS–H + M*, and 3) M* → M + $h\nu$. In this case, only a single electron has been transferred during hydrogen condensation.

Perhaps, it is through the examination of linewidths that the most interesting conclusions can be reached. The emission lines of Hα, Ca, and Mg from spicules are very broad, suggesting a strong interaction between CHS and the ejected atoms, in association with ejection and light emission [234–236]. In contrast, spicule emission linewidths from Hβ, Hγ, Hϵ, the D3 line from He, and the neutral line from oxygen are all sharp [234]. One could surmise that the interaction between these species and condensed hydrogen structures are weaker upon ejection.

It is reasonable to conclude that the hydrides play an important role in facilitating condensation within the chromosphere [59]. Hydrides enable the delivery of hydrogen in a systematic manner and, most importantly, either one or two electrons, depending on the electron densities present on the local level. Such an elegant mechanism to account for the prevalence of CaII and MgII in the chromosphere cannot be achieved by other models. Moreover, unlike the LMH model, the gaseous models take no advantage of the chemical species known to exist in the solar atmosphere.

3.6 Helium Emission #14

The analysis of helium emission in the chromosphere may well provide the most fascinating adventure with regard to the spectroscopic lines of evidence.‡ This stands as fitting tribute to helium [47], as it was first observed to exist on the Sun [237, 238]. These seminal discoveries exploited the presence of helium within prominences and the disturbed chromosphere [239, 240]. Astronomers would come to view solar helium as extremely abundant [241, 242], but these con-

Here is a brief list of interesting ions and the ionization energies required for their production: HII = 13.6 eV; HeII = 24.6 eV; HeIII = 54.4 eV; MgII = 7.6 eV; MgIII = 15.0 eV; CaII = 6.1 eV; CaIII = 11.8 eV and FeXIV = 361 eV [233]. In this respect, note how the first ionized form of helium, HeII, requires 24.6 eV for its production. The generation of many triplet forms of orthohelium HeI will demand energies of ∼20 eV. To remove two electrons from calcium yielding CaIII (the stable Ca^{+2} ion) only requires 11.8 eV. As a result, how can the gas models account for the presence of CaII lines at high altitude on the Sun (5-10,000 km), when this ion only requires 6.1 eV for production? If such powerful HeII and HeI* can be observed, why is CaIII, which requires only 11.8 eV for its generation and has the inert gas, [Ar], configuration, not the preferred form of calcium? This provides a powerful clue that the presence (or absence) of an individual ion on the Sun is related to chemistry and not to temperature.

†As will be seen in §3.8, it is envisioned that the corona of the Sun is harvesting electrons.

‡This proof was first presented as the 32nd line of evidence [61].

clusions have been challenged and may need to be revisited [47, 48, 61]. There is considerable reason to conclude that the solar body is actively ejecting He from its interior [47, 48].

Though helium can be found in spicules [193] and prominences, it is difficult to observe on the solar disk. It can be readily visualized in the chromosphere where the spatial extent of the 30.4 nm HeII emission lines can greatly exceed those from Hα (see the wonderful Fig. 1 in [243]). With increased solar activity, helium emission can become pronounced in the solar atmosphere (see Fig. 15 and [244]).

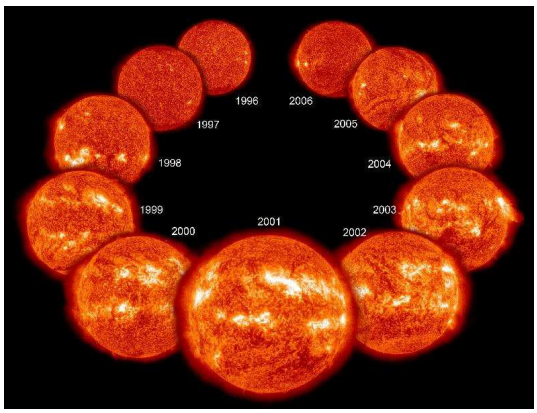

Fig. 15: Image of consecutive years in the solar cycle taken in the HeII line at 30.4 nm. NASA describes this image as follows, *"An EIT image in the 304 Angstrom wavelength of extreme UV light from each year of nearly an entire solar cycle"*. Courtesy of SOHO/[EIT] consortium. SOHO is a project of international cooperation between ESA and NASA. (http://sohowww.nascom.nasa.gov/gallery/images/cycle002.html — Accessed on 9/20/2013).

In the chromosphere, the helium which gives rise to emission lines can possess both of its electrons (HeI) or lose an electron to produce an ion (HeII). HeII resembles the hydrogen atom in its electronic configuration. However, the situation concerning HeI can be more complex. When this species exists in the ground state, both of its electrons lie in the 1S orbital (N=1) with their spins antiparallel, as dictated by Pauli's exclusion principle. In the excited state (i.e. 1 electron in the N=1 shell, and the second electron in any of the N>1 shells), helium can exist either as a singlet (parahelium — spins remaining antiparallel to one another) or as a triplet (orthohelium — spins assume a parallel configuration). Interestingly, the line emissions from the triplet states of orthohelium can be quite strong on the limb of the Sun.

For instance, a well-known triplet HeI transition occurs at 1083 nm (10830Å) which is barely visible on the disk, but it is nearly as intense as Hα on the limb [245, p. 199–200]. At the same time, the HeI triplet D3 line at 588 nm can be enhanced 20 fold when visualization moves from the disk to the limb [245, p. 199-200].*

During the eclipse of March 29, 2006, the triplet D3 line was carefully examined. It appeared to have a binodal altitude distribution with a small maximum at ~250 km and a stronger maximum between 1300-1800 km (see Fig. 6 in [244]). This bimodal distribution was not always observed (see Fig. 7 in [244]). But generally, the D3 line is most intense at an altitude of ~2,000 km, with an emission width of approximately 1,600 km. The triplet D3 lines show no emission near the photosphere.

Within the context of gaseous models, it is extremely difficult to account for the presence of excited HeI triplet states in the chromosphere. Helium requires ~20 eV† to raise an electron from the N=1 shell to the N=2 shell. How can excitation temperatures in excess of 200,000 K be associated with a chromosphere displaying apparent temperatures of 5,000-10,000K, values not much greater than those existing on the photosphere?

Therefore, since proponents of gaseous models are unable to easily account for the powerful D3 line emission, they have no choice but to state that helium is being excited by coronal radiation which has descended into the chromosphere [244, 246]. In a sense, helium must be *'selectively heated'* by the corona. These proposals strongly suggest that the gaseous models are inadequate. It is not reasonable to advance that an element can be selectively excited by coronal radiation, and this over its many triplet states. At the extreme, these schemes would imply that coronal photons could strip away all electrons from chromospheric atoms. Yet, even lines from neutral atoms are observed.‡

On the other hand, helium emissions can be easily understood in the LMH model [35, 36, 39], if attention is turned toward condensation reactions believed to occur within the chromosphere (see §3.4, §3.5 and [59, 61]).

In this respect, it must be recognized that the famous helium hydride cation (HeH^+) *"is ubiquitous in discharges containing hydrogen and helium"* [247].

First discovered in 1925 [248], HeH^+ has been extensively studied [249, 250] and thought to play a key role in certain astrophysical settings [251–253]. In the laboratory, its spectral lines were first observed by Wolfgang Ketterle (Nobel Prize, Physics, 2001) [254, 255]. The author has previously noted, *"Although it exists only in the gas phase, its Brønsted acidity should be extremely powerful. As a result, the hydrogen hydride cation should have a strong tendency to donate a proton, without the concerted transfer of an electron"* [61].

Turning to Fig. 16, it appears that the action of the helium hydride cation, HeH^+, can lead to a wide array of reactions within the chromosphere. These processes are initiated with

*Lines from neutral helium can be enhanced 50 fold on the limb relative to the disk [245, p. 199-200].

†1 eV = 11,600 K ; 20 eV = 232,000 K.

‡Selective excitation was also used to account for the emission lines from molecular hydrogen [220]. But it is more likely that these reflect the delivery of a hydrogen cluster (see §3.4.1) with H_2^* rather than H^* expulsion.

its transfer to condensed hydrogen structures, CHS, believed to be be forming (see §2.3.6, §3.4, §3.5, §3.7, §5.4, §5.6, §6.6) in this region of the solar atmosphere. As was the case with hydrogen (§3.4) and elemental (§3.5) emission lines, everything hinges on the careful consideration of condensation.

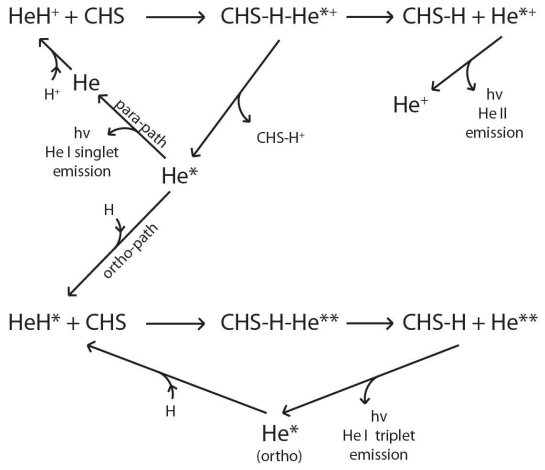

Fig. 16: Schematic representation of possible pathways involved when the helium hydride ion, HeH^+, or the excited helium hydride molecule, HeH^*, react with condensed hydrogen structures, CHS, in the chromosphere of the Sun. The pathways presented can account for all emission lines observed from He I and He II. Note in this scheme that excited helium, He^*, is being produced initially through the interaction of HeH^+ with CHS. This excited helium, He^*, if it assumes the triplet state (orthohelium — electrons in the same orientation: spin up/up or down/down), will become trapped in excited state. This triplet helium can then be used repeatedly, in cyclic fashion, to condense hydrogen atoms onto chromospheric structures, CHS (as shown in the lower half of the figure). Alternatively, if excited helium He^* is initially produced in the singlet state (parahelium — electrons in different orientation: spin up/down), emission can immediately occur generating the singlet lines from He I. This scheme accounts for the strong triplet He I transition at 10830 Å observed in the flash spectrum of the chromosphere. Unlike the situation in the gas models, random collisional or photon excitations are not invoked to excite the helium atoms. De-excitation processes would also be absent, helping to ensure the buildup of triplet state orthohelium in this model. This figure, along with its legend, was previously published in [61].

First, HeH^+ and CHS react to form an activated complex: $CHS + HeH^+ \rightarrow CHS\text{-}H\text{-}He^{+*}$. If the expulsion of an excited helium ion (He^{+*}) follows, full transfer of a proton and an electron to CHS will have occurred (top line in Fig. 16). The resulting He^{+*} would be able to relax back to a lower energy state through emission, leading to the well known He II lines in the chromosphere (top right in Fig. 16).

Alternatively, when HeH^+ reacts with CHS, the expulsion of an excited helium atom (He^*) could follow (see Fig. 16) involving the transfer of a proton — but no electron — to the CHS. As a strong Brønsted acid, HeH^+ should permit these reactions (namely: $CHS\text{–}HHe^{+*} \rightarrow CHS\text{–}H^+ + He^*$). Expulsion of an activated helium atom (He^*) can lead to two conditions, depending on whether the electrons within this species are antiparallel (parahelium) or parallel (orthohelium). Within helium, the excited electron is allowed by selection rules to return to the ground state, if and only if, its spin is opposed to that of the ground state electron. As a result, only parahelium can relax back to the ground state: $He^* \rightarrow He + h\nu$. This leads to the HeI lines from singlet helium.

As for the excited orthohelium, it is unable to relax, as its two electrons have the same spin (either both spin up or both spin down). *Trapped* in the excited state, this species can at once react with hydrogen, forming the excited helium hydride molecule, which, like the helium hydride cation, is known to exist [256, 257]: $He^* + H \rightarrow HeH^*$.

Excited helium hydride can react with CHS in the chromosphere, but now resulting in a doubly activated complex: $CHS + HeH^* \rightarrow CHS\text{-}H\text{-}He^{**}$, wherein one electron remains in the ground state and the other electron is promoted beyond the 2S shell.* To relax, the doubly excited He^{**} atom, must permit an electron currently in the 2P or higher orbital, to return to the 2S or 2P orbitals.

The helium D_3 line would be produced by a $3^3D \rightarrow 2^3P$ transition [245, p. 95]. The $2^3P \rightarrow 2^3S$ transition is associated with the strong triplet He I line at 10830 Å [245, p. 95]. Alternatively, a $3^3P \rightarrow 2^3S$ transition produces the triplet He I line at 3890 Å [245, p. 95].

Importantly, since excited orthohelium cannot fully relax back to the ground state, it remains available to recondense with atomic hydrogen in the chromosphere. This results in its continual availability in the harvest of hydrogen. A cyclic process has been created using orthohelium (He^*). The priming of this cycle had required but a single instance where hydrogen was transferred to CHS by HeH^+, without the complementary transfer of an electron (top line in Fig. 16).† In this manner, much like what occurred in the case of molecular hydrogen (§3.4) and the metal hydrides (§3.5), the body of the Sun has been permitted to recapture atomic hydrogen lost to its atmosphere. It does not simply lose these atoms without any hope of recovery [59, 61, 62].

Within the LMH model, the prominence of the helium triplet lines can be elegantly explained. They result from the systematic excitation of helium, first delivered to condensed hydrogen structures by the helium hydride cation (HeH^+), a well-known molecule [247–254] and strong Brønsted acid. The generation of triplet state excited helium can be explained in a systematic fashion and does not require unrealistic temperatures in the corona. It is not an incidental artifact produced by improbably selective excitations generated using

*The possibility that He^{**} could have no electrons in the ground state is not considered.

†The production of Ca II emission lines from CaH had resulted in the transfer of two electrons per hydrogen atom (see §3.5). This can help keep charge neutrality in condensation reactions involving HeH^+.

coronal photons. Organized chemical reactions govern the behavior of helium in the Sun, not random events.

3.7 Fraunhofer Absorption #15

When examined under high spectral resolution, the visible spectrum of the Sun is punctuated by numerous absorption lines, which appear as dark streaks against a brighter background.* These lines were first observed by William Hyde Wollaston in 1802 [258]. They would eventually become known as *Fraunhofer lines* after the German scientist who most ably described their presence [259]. Fraunhofer lines can be produced by many different elements. They manifest the absorption of photospheric light by electrons, contained within gaseous atomic or ionic species above the photosphere, which are being promoted from a lower to a higher energy level.

In 1862, Kirchhoff was the first to argue that the Fraunhofer lines provided evidence for a condensed solar body, *"In order to explain the occurrence of the dark lines in the solar spectrum, we must assume that the solar atmosphere incloses a luminous nucleus, producing a continuous spectrum, the brightness of which exceeds a certain limit. The most probable supposition which can be made respecting the Sun's constitution is, that it consists of a solid or liquid nucleus, heated to a temperature of the brightest whiteness, surrounded by an atmosphere of somewhat lower temperature."* [190, p. 23].

Amongst the most prominent of the Fraunhofer lines are those associated with the absorption of photospheric light by the hydrogen atoms. The preeminent Fraunhofer lines are generated by the Balmer series. These lines are produced when an excited hydrogen electron (N=2) absorbs sufficient energy to be promoted to yet higher levels ($H\alpha$ $N=2\rightarrow N=3$ 656.3 nm; $H\beta$ $N=2\rightarrow N=4$ 486.1 nm; $H\gamma$ $N=2\rightarrow N=5$ 434.1 nm; $H\delta$ $N=2\rightarrow N=6$ 410.2 nm; etc). They can be readily produced in the laboratory by placing hydrogen gas in front of a continuous light source.

In 1925, Albrecht Unsöld reported that the solar Fraunhofer lines associated with hydrogen did not decrease as expected [260]. He noted intensities across the Balmer series ($H_\alpha = 1$; $H_\beta = 0.73$; $H_\gamma = 0.91$; $H_\delta = 1$) which where highly distorted compared to those expected in a hydrogen gas, as predicted using quantum mechanical considerations ($H_\alpha = 1$; $H_\beta = 0.19$; $H_\gamma = 0.07$; $H_\delta = 0.03$) [260].

Hydrogen lines were known to be extremely broad from the days of Henry Norris Russell and Donald H. Menzel, who had observed them in association with solar abundance [87] and chromospheric studies [205], respectively. Commenting on the strength of the hydrogen Balmer series, Henry Norris Russell would write, *"It must further be born in mind that even at solar temperatures the great majority of the atoms of any given kind, whether ionized or neutral, will be in the state of lowest energy... One non-metal, however, presents a real and glaring exception to the general rule. The hydrogen lines of the Balmer series, and, as Babcock has recently shown, of the Paschen series as well, are very strong in the Sun, though the energy required to put an atom into condition to absorb these series is, respectively, 10.16 and 12.04 volts — higher than for any other solar absorption lines. The obvious explanation — that hydrogen is far more abundant than the other elements — appears to be the only one"* [87, p. 21–22].

In the photospheric spectrum, the hydrogen absorption lines are so intense that the observer can readily garner data from the Lyman (N=1 → N=2 or higher), Balmer (N=2 → N=3 or higher), Paschen (N=3 → N=4 or higher), and Brackett (N=4 → N=6 or higher) series [87, 205, 260–264].

The central questions are three fold: 1) Why are the hydrogen lines broad? 2) Why does hydrogen exist in excited state as reflected by the Balmer, Paschen, and Brackett lines? and 3) Why is the normal quantum mechanical distribution of the Balmer series distorted as first reported by Unsöld [260]?

In the gaseous models, different layers of the solar atmosphere have to be invoked to account for the simultaneous presence of Lyman, Balmer, Paschen and Brackett line profiles in the solar spectrum [261–264]. Once again, as when addressing limb darkening (see §2.3.2), the models have recourse to optical depth [261–264]. These approaches fail to adequately account for the production of the excited hydrogen absorption.

As noted in §3.4, in the setting of the LMH model, excited hydrogen atoms can be produced through condensation reactions occurring in the solar chromosphere. These atoms could be immediately available for the absorption of photons arising from photospheric emission. Hence, condensation reactions provide an indirect mechanism to support the generation of many hydrogen Fraunhofer line. Since these lines are being produced in close proximity to condensed matter, it is reasonable to conclude that their linewidths are determined by their interaction with such materials and not from optical depth and Stark mechanisms (see §3.4). This may help to explain why the intensity of the Balmer lines, as first reported by Unsöld [260], do not vary as expected in gases from quantum mechanical considerations. Unsöld's findings [260] strongly suggest that the population of excited hydrogen atoms is being distorted by forces not known to exist within gases. Once again, this calls attention to condensed matter.

3.8 Coronal Emission #16

As was discussed in §2.3.7, the K-corona is the site of continuous emission which reddens slightly with altitude, but whose general appearance closely resembles the photospheric spectrum [57].† This leads to the conclusion that condensed matter must be present within this region of the Sun [57]. Still, the nature of the corona is more complicated, as the same region which gives rise to condensed matter in the K-

*This proof was first presented as the sixteenth line of evidence [47, 59].

†This proof was first presented as the 31st line of evidence [60, 62].

corona is also responsible for the production of numerous emission lines from highly ionized elements (e.g. FeXII-FeXXV [192]) in the E-corona [60].*

When examined in light of the gaseous solar models, the production of highly ionized species requires temperatures in the million of degrees [192]. Temperatures as high as 30 MK have been inferred to exist in the corona [192, p. 26], even if the solar core has a value of only 16 MK [13, p. 9]. Flares have been associated with temperatures reaching 10^8 K [273], and radio sampling has called for values between 10^8 and 10^{10} K [245, p. 128].

Given the temperatures inferred in attempting to explain the presence of highly ionized atoms in the K-corona, proponents of the gaseous models deny that this region can be comprised of condensed matter. Harold Zirin summarizes the situation best, *"... there is something erroneous in our basic concept of how ionization takes place"* [245, p. 183].

Rather than cause a dismissal of condensed matter, such extreme temperature requirements should lead to the realization that the gaseous models are fundamentally unsound [62]. It is not reasonable to assume that the corona harbors temperatures which exceed those found in the core. Furthermore, to arrive at these extreme values, the corona must somehow be heated. The *"zoo"* [148, p. 278] of possible heating mechanisms is substantial [148, p. 239–251]. According to E.R. Priest, the hypothesized mechanisms are fundamentally magnetic in nature as *"all the other possible sources are completely inadequate"* [273]. The problem for gaseous models can be found in the realization that their only means of producing highly ionized atoms must involve violent bombardment and the removal of electrons to infinity. These schemes demand impossible temperatures.†

It is more reasonable to postulate that elements within the corona are being stripped of their electrons when they come into contact with condensed matter. The production of highly ionized atoms involves electron affinity, not temperature. The belief that the corona is a region characterized by extremely elevated temperatures is erroneous. The cool K-coronal spectrum is genuine. The associated photons are directly produced by the corona itself, not by the photosphere (see §2.3.7).

Moreover, condensed matter can have tremendous electron affinities. This is readily apparent to anyone studying lightning on Earth. Thunderhead clouds have been associated with the generation of 100 keV X-rays [274, p. 493-495], but no-one would argue that the atmosphere of the Earth sustains temperatures of 10^9 K. Lightning can form *"above volcanoes, in sandstorms, and nuclear explosions"* [274, p. 67]. It represents the longest standing example of the power of electron affinity, as electrons are transferred from condensed matter in the clouds to the Earth's surface, or vise versa [274–276].

Metallic hydrogen should exist in the K-corona, as Type-I material has been ejected into this region (see §2.3.8) by activity on the photosphere [58]. Electrical conductivity in this region is thought to be very high [277, p. 174]. Thus, the production of highly ionized elements can be explained if gaseous atoms come into contact with this condensed matter. For example, iron (Fe) could interact with metallic hydrogen (MH) forming an activated complex: MH + Fe $\rightarrow$ MH–Fe*. Excited Fe could then be ejected with an accompanying transfer of electrons to metallic hydrogen: MH–Fe* $\rightarrow$ MH–$n\bar{e}$ + Fe^{+n*}. The emission lines observed in the corona are then produced when the excited iron relaxes back to the ground state through photon emission, Fe^{+n*} $\rightarrow$ Fe^{+n} + $h\nu$. Depending on the local electron affinity of the condensed metallic hydrogen, the number of electrons transferred, n, could range from single digits to ~25 [192] in the case of iron.‡

The scheme formulated with iron can be extended to all the other elements,§ resulting in the production of all coronal emission lines. The governing force in each case would be the electron affinity of metallic hydrogen which may increase with altitude. Highly ionized species are not produced through the summation of multiple electron ejecting bombardments. Rather, multiple electrons are being stripped simultaneously, in single action, by transfer to condensed matter. In this manner, the *electron starved* corona becomes endowed with function, *the harvesting of electrons from elements in the solar atmosphere, thereby helping to maintain the neutrality of the solar body* [60].

In this sense, the chromosphere and corona have complimentary action. The chromosphere harvests hydrogen atoms and protons. The corona harvests electrons.¶

As for the transition zone (see Fig. 1.1 in [192]), it does not exist. This region was created by the gaseous models in order to permit a rapid transition in apparent temperatures between the cool chromosphere and hot corona (see [62] for a complete discussion). In the metallic hydrogen model, the apparent temperatures in both of these regions are cool, there-

*The story which accompanies the mystical element coronium (or FeXIV) in the corona and its discovery by the likes of Harkness, Young, Grotian, and Edlén [151–153] has been recalled [265–268]. Wonderful images of the corona have recently been produced from highly ionized iron (e.g. FeX-FeXIV) [269–272].

†It will be noted in §5.5, that the gaseous solar models infer widely varying temperatures within the *same* regions of the corona when analyzing coronal loops (see Fig. 22). How could it be possible to sustain vastly differing values in the *same* region of the solar atmosphere? These findings are indicative that we are not sampling temperature, but rather substructures with distinct electron affinities. These substructures take advantage of a wide array of species to transfer electrons. Evidence for such a solution can be found in Fig. 1.10 of [192] which describes flare substructure and the associated variations in emitting species (arcade emitting in FeXII — spine emitting in FeXXIV and Ca XVII).

‡In this regard, it is important to note that most of the ions present in the *"XUV spectrum are principally those with one or two valence electrons"* [245, p. 173]. This observation is highly suggestive that systematic processes are taking place, not random bombardments.

§A least one electron must remain for line emission.

¶While the corona is primarily composed of metallic hydrogen, as will be seen in §5.4, it can provide a framework to allow for the condensation of hydrogen in non-metallic form.

fore a transition zone serves no purpose [62]. The changes in atomic and ionic compositions observed in the solar atmosphere can be accounted for by 1) the varying ability of molecular species to deliver hydrogen and protons to condensed hydrogen structures in the chromosphere as a function of altitude, and 2) to changes in the electron affinity of metallic hydrogen in the corona.

This scenario resolves, at long last, the apparent violation of the Second Law of Thermodynamics which existed in the gaseous model of the Sun. It is not realistic that the center of the Sun exists at 16 MK [13, p. 9], the photosphere at 6,000 K, and the corona at millions of degrees. A solution, of course, would involve the recognition that most of the energy of the photosphere is maintained in its convection currents and conduction bands [37], not in the vibrational modes responsible for its thermal spectrum and associated apparent temperature. But now, the situation is further clarified. The corona is not being heated — it is cool. No violation of the Second Law of Thermodynamics exists, even if photospheric convection and conduction are not considered.

4 Structural Lines of Evidence

The structural lines of evidence are perhaps the most physically evident to address, as they require only elementary mechanical principles to understand.

4.1 Solar Collapse #17

Should stars truly be of gaseous origin, then they are confronted with the problem of solar collapse.* Somehow, they must prevent the forces of gravity from causing the entire structure to implode upon itself.

Arthur Eddington believed that stellar collapse could be prevented by radiation pressure [9]. Photons could transfer their momentum to stellar particles and thereby support structure. These ideas depend on the existence of radiation within objects, a proposal which is counter to all laboratory understanding of heat transfer. Conduction and convection are responsible for the transfer of energy within objects [70]. It is only if one wishes to view the Sun as an assembly of separate objects that radiation can be invoked.

Eventually, the concept that the Sun was supported exclusively by radiation pressure was abandoned. Radiation pressure became primarily reserved for super-massive stars [13, p. 180-186]. Solar collapse was prevented using *'electron gas pressure'* [13, p. 132], with radiation pressure contributing little to the solution [13, p. 212].

But the idea that *'electron gas pressure'* can prevent a star from collapsing is not reasonable [3, 35, 43, 48]. The generation of gas pressure (see Fig. 17) requires the existence of true surfaces, and none can exist within a gaseous Sun.† When a particle travels towards the solar interior, it can simply undergo an elastic collision, propelling a stationary particle beneath it even further towards the core. Without a surface, no net force can be generated to reverse this process: the gaseous Sun is destined to collapse under the effect of its own gravity [48].

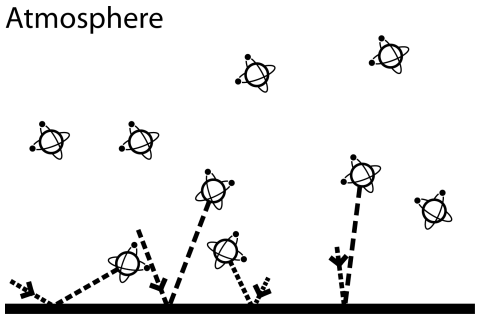

Fig. 17: Schematic representation of the generation of gas pressure. As particles travel towards a real surface, they eventually undergo a change in direction resulting in the creation of a net upwards force.

Donald Clayton, a proponent of the gaseous models, describes the situation as follows, "*The microscopic source of pressure in a perfect gas is particle bombardment. The reflection (or absorption) of these particles from a real (or imagined) surface in the gas results in a transfer of momentum to that surface. By Newton's second law* ($F = dp/dt$)*, that momentum transfer exerts a force on the surface. The average force per unit area is called the pressure. It is the same mechanical quantity appearing in the statement that the quantity of work performed by the infinitesimal expansion of a contained gas is* $dW = PdV$*. In thermal equilibrium in stellar interiors, the angular distribution of particle momenta is isotropic; i.e., particles are moving with equal probabilities in all directions. When reflected from a surface, those moving normal to the surface will transfer larger amounts of momentum than those that glance off at grazing angles*" [14, p. 79]. The problem is that real surfaces do not exist within gaseous stars and *'imagined'* surfaces are unable to be involved in a *real* change in momentum. *'Electron gas pressure'* cannot prevent solar collapse.

Unlike the scenario faced by Eddington with respect to solar collapse, James Jeans had argued that liquid stars were immune to these complications, "*And mathematical analysis shews that if the centre of a star is either liquid, or partially so, there is no danger of collapse; the liquid center provides so firm a basis for the star as to render collapse impossible*" [278, p. 287]. By their very nature, liquids are essentially incompressible. Therefore, liquid stars are self-supporting and a LMH Sun faces no danger of collapse.

*This proof was first presented as the third line of evidence [3, 35, 43, 48].

†Conversely, the extended nature of our atmosphere is being maintained through gas pressure precisely because our planet possesses a real surface. When gas particles strike the Earth's surface, they undergo an immediate change in direction with upward directed velocities. Without the presence of a true surface, a net change in particle velocity cannot occur.

4.2 Density #18

Hot gases do not *self-assemble*.* Rather, they are well-known to rapidly diffuse, filling the volume in which they are contained. As a result, hot gaseous *'objects'* should be tenuous in nature, with extremely low densities. In this respect, hot gases offer little evidence that they can ever meet the requirements for building stars.

In an apparent contradiction to the densities expected in gaseous *'objects'*, the solar body has a substantial average density on the order of 1.4 g/cm^3 [279]. In gaseous models, the Sun is believed to have a density approaching 150 g/cm^3 in its core, but only $\sim 10^{-7}$ g/cm^3 at the level of the photosphere [148]. In this way, a gaseous star can be calculated with an average density of 1.4 g/cm^3. But gaseous models would be in a much stronger position if the average density of the Sun was consistent with that in a sparse gas, i.e. $\sim 10^{-4}$ g/cm^3, for instance. It is also concerning that the average density of the Sun is very much coincident with that observed in the outer planets, even though these objects have much smaller total masses.† The giant planets are no longer believed to be fully gaseous, but rather composed of metallic hydrogen [93–95], suggestions which are contrary to the existence of a gaseous Sun.

The Sun has a density entirely consistent with condensed matter. If the solar body is assembled from metallic hydrogen [35, 39], it is reasonable to presume that it has a somewhat uniform distribution throughout its interior.‡ This would be in keeping with the known, essentially incompressible, nature of liquids.

4.3 Radius #19

Within gaseous models, the Sun's surface cannot be real and remains the product of optical illusions [2, 4, 51].§ These conjectures were initially contrived by the French astronomer, Hervé Faye. In 1865, Faye [280] had proposed that the Sun was gaseous [2, 4] and would write, *"This limit is in any case only apparent: the general milieu where the photosphere is incessantly forming surpasses without doubt, more or less, the highest crests or summits of the incandescent clouds, but we do not know the effective limit; the only thing that one is permitted to affirm, is that these invisible layers, to which the name atmosphere does not seem to me applicable, would not be able to attain a height of 3', the excess of the perihelion distance of the great comet of 1843 on the radius of the photosphere"* [280]. With those words, the Sun lost its true surface. Everything was only *'apparent'* (see §1). Real dimensions, like diameter or radius, no longer held any validity. Nonetheless, Father Secchi considered the dimensions of the Sun to be a question of significant observational importance, despite problems related to their accurate measure [1, p. 200–202, V. I].

Today, the radius of the Sun (~696,342±65 km) continues to be measured [51] and with tremendous accuracy — errors on the order of one part in 10,000 or even 2 parts in 100,000 (see [281] for a table). Such accurate measurements of spatial dimensions typify condensed matter and can never characterize a gaseous object.¶ They serve as powerful evidence that the Sun cannot be a gas, but must be composed of condensed matter.

The situation relative to solar dimensions is further complicated by the realization that the solar diameter may well be variable [282]. Investigations along these lines are only quietly pursued [283], as the gas models are unable to easily address brief fluctuations in solar dimensions. The stability of gaseous stars depends on hydrostatic equilibrium and relies on a perfect mechanical and thermal balance [13, p. 6–67]. Failing to maintain equilibrium, gaseous stars would cease to exist.

Conversely, fluctuating solar dimensions can be readily addressed by a liquid metallic hydrogen Sun, since this entity enables localized liquid/gas (or solid/gas) transitions in its interior (see [48, 51, 52] and §5.1).

4.4 Oblateness #20

James Jeans regarded the high prevalence of binaries as one of the strongest lines of evidence that the stars were liquids [27, 28].‖ Indeed, it could be stated that most of his thesis rested upon this observation. As a spinning star became oblate, it eventually split into two distinct parts [27, 28]. Oblateness can be considered as a sign of internal cohesive forces within an object and these are absent within a gaseous star. As a result, any oblateness constitutes a solid line of evidence that a rotating mass is comprised of condensed matter.

The physics of rotating fluid masses has occupied some of the greatest minds in science, including Newton, Maclaurin, Jacobi, Meyer, Liouville, Dirichlet, Dedekind, Riemann, Poincaré, Cartan, Roche, and Darwin [3]. The problem also captivated Chandrashekhar (Nobel Prize, Physics, 1983) for nine years of his life [284].

Modern studies placed the oblateness of the Sun at 8.77×10^{-6} [287]. Though the Sun appears almost perfectly

*This proof was first presented as the fourth line of evidence [35, 36].

†The Earth has a density of 5.5 g/cm^3; Jupiter 1.326 g/cm^3; Saturn 0.687 g/cm^3; Neptune 1.638 g/cm^3; Uranus 1.271 g/cm^3 [279].

‡Setsuo Ichimaru had assumed, based on the gaseous models, that the core of the Sun had a density of 150 g/cm^3 when he considered that it could be composed of metallic hydrogen [97–99]. He did not address the composition of the solar body or atmosphere.

§This proof was first presented as the 21st line of evidence [51].

¶As a point of reference relative to the accuracy of measurements, machinists typically work to tolerances of a few thousands of an inch. According to a young machinist (Luke Ball, Boggs and Associates, Columbus, Ohio), a *"standard dial caliper is accurate to ± 0.001", and a micrometer provides greater accuracy to ± 0.0001". The Mitutoyo metrology company was founded in 1934, and they produce a digital high-accuracy sub-micron micrometer that is accurate to .00002."*

‖This proof was first presented as the eighth line of evidence [3, 35, 36, 50].

round, it is actually oblate [50].* To explain this behavior, astrophysicists invoked that the Sun possessed a constant solar density as a function of radial position [287]. This proposal is in direct conflict with the gaseous solar models [13, 14] which conclude that most of the solar mass remains within the central core. An essentially constant internal density is precisely what would be required within the context of a liquid metallic Sun [35, 39].

At present, helioseismic measurements (see §6) indicate that the degree of solar oblateness may be slightly smaller [288, 289], but the general feature remains. The degree of solar oblateness may well vary with the solar cycle [290]. As was the case for variations in solar radius (§4.3), these changes pose difficulties for the gaseous models. That the Sun is slightly oblate provides excellent evidence for internal cohesive forces, as seen in condensed matter.

4.5 Surface Imaging #21

With the advent of the 1-m Swedish Solar Telescope (SST), the solar surface has been imaged with unprecedented resolution [100, 291].† This resolution will increase dramatically in a few years when the construction of the Advanced Technology Solar Telescope is completed in Hawaii [104].

Using the SST, scientists report, *"In these pictures we see the Sun's surface at a low, slanting angle, affording a three-dimensional look at solar hills, valleys, and canyons"* [291]. . . *"A notable feature in our best images of sunspots is that many penumbral filaments, which are isolated from the bulk of the penumbra and surrounded by dark umbra, show dark cores"*. . . *"Inspection of our images shows numerous varieties of other very thin dark lines in magnetic regions"* . . . *"'hairs' that are seemingly emanating from pores into the closest neighbouring granules, 'canals' in the granulation near spots and pores, and running dark streaks crossing penumbral filaments diagonally"* [100].

Since antiquity, solar observers have been fascinated with structure on the surface of the Sun. Now, as telescopic resolution continues to increase, they are documenting, *almost in 3D*, the existence of structure on the solar surface with increased certainty. They resort to words like *'hills'*, *'valleys'*, and *'canyons'* to describe the surface of the Sun and they focus increasingly on substructures, like the dark cores of the penumbra. How can this structural detail be compatible with gases? Structure remains a property of condensed matter and gases can support none. Moreover, if the solar surface is but an *'illusion'*, what point can there be in documenting the nature of these structures? But the problem is even more vexing for the gaseous models, as films are currently being taken of the Sun in high resolution (see Supplementary Materials for [100] on the Nature website), and our *'illusions'* are *behaving* as condensed matter (see §5.1) [292, 293].

Father Secchi, perhaps the most able solar observer of the 19th century, drew with painstaking attention numerous details on the solar surface which he viewed as real [1]. He emphasized that *"there is thus no illusion to worry about, the phenomena that we have just exposed to the reader are not simple optical findings, but objects which really exist, faithfully represented to our eyes using instruments employed to observe them"* [1, p. 35–36, V. II]. The authors of the wonderful SST Nature paper [100] seem to discard illusions, *"We are, however, confident that the dark cores shown here are real"* [100]. Nonetheless, they maintain the language associated with the gaseous models, *"A dark-cored filament could be produced by an optically thin cylindrical tube with hot walls—perhaps a magnetic flux tube heated on the surface by the dissipation of electrical currents"* [100].

Commenting on [100] in light of accepted theory, John H. Thomas states, *"Computer simulations of photospheric magnetoconvection show very small structures, but the simulations have not yet achieved sufficient resolution to determine the limiting size. The horizontal mean free path — in other words, the average distance traveled without interacting — of a photon in the solar photosphere is about 50 km, and so this might be expected to be the smallest observable length scale, because of the smoothing effect of radiative energy transfer. But sophisticated radiative-transfer calculations show that fine structures as small as a few kilometers should in principle be directly observable"* [294].

The problem for the gas models rests in their prediction that the photosphere has a density ($\sim 10^{-7}$ g/cm^3 [148]) which is 10,000 times lower than that of the Earth's atmosphere at sea level — surpassing some of the best vacuums on Earth. Structure cannot be claimed to exist in a vacuum and has never been demonstrated to be associated with the equations of radiation transfer (see [292, 294] and references therein). It is inherently a property of condensed matter, without any need for internal photons. As a result, modeling associated with the analysis of structural entities on the solar surface, which is fundamentally based on ideas of a gaseous Sun [292, 294], are unlikely to be of any lasting value with respect to understanding the complexities of the photosphere. The most elegant solution rests in accepting that these structures are real and comprised of condensed matter.

4.6 Coronal Holes/Rotation #22

Coronal holes (see Fig. 18) are believed to be regions of low-density plasma that open freely into interplanetary space [52,

*As a point of interest, the Southern star Achernar, has a tremendous oblateness which approaches 1.5 [285]. This value cannot be explained using the standard gaseous models wherein most of a star's mass is restricted to the core. As such, scientists have sought to find alternative means to account for this oblateness [286].

†This proof was first presented as the eleventh line of evidence [4, 35, 36, 42]. Solar surface imaging can include frequencies outside visible light. It continues to reveal the presence of new structures, not described in §2. These, and those to come, are included herein as a separate line of evidence as solar surface imaging exposes more structural complexity and temporal evolution.

295, 296].* They are associated with the presence of fast solar winds (see §5.8).

When the Sun becomes active, coronal holes can appear anywhere on the solar surface [52, 295, 296]. In contrast, when it is quiet, coronal holes are viewed as *'anchored'* onto the polar regions of the solar surface [297, p. 10]. This *'anchoring'* constitutes a powerful sign that the Sun is comprised of condensed matter, as this behavior directly implies both long-term structure within the corona and the existence of a true solar surface. *'Anchoring'* requires two distinct regions in the Sun which cooperate with each other to produce *structural restriction*.

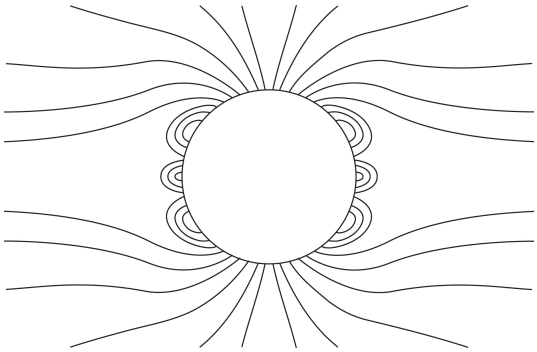

Fig. 18: Schematic representation of coronal holes over the polar caps of a quiet Sun. This figure is an adaptation based on Fig. 2 in [295]. Along with its legend, it was previously published in [52].

The corona possesses *"... a radially rigid rotation of 27.5 days synodic period from 2.5 $R_\odot$ to >15$R_\odot$"* [277, p. 116] as established by the LASCO instrument aboard the SOHO satellite [298]. Rigid rotation of the entire corona strongly suggests that the solar body and the corona possess condensed matter.

Coronal material† contains magnetic fields lines which, in turn, are anchored at the level of the photosphere [62]. *'Anchoring'*, once again, requires structure both within the solar body and within the solar atmosphere. The condensed nature of the corona and coronal structures has already been discussed in §2.3.7, §2.3.8, and §3.8. It will be treated once again in §5.5, and §6.6. The relevant structure of the solar interior will be discussed in §5.1. The presence of *'anchoring'* within coronal holes and the rigid rotation of the corona is best explained by condensed matter.

4.7 Chromospheric Extent #23

Eddington recognized the great spatial extent of the chromosphere and pondered on how this material was supported [9, p. 362].‡ At the time, he knew that chromospheric emission lines (see §3.4, §3.5, and §3.6) could extend up to 14,000 km [9, p. 362]. For Eddington, the answer to chromosphere chromospheric extent rested upon radiation pressure, but the solution would prove insufficient [62].

Bhatnagar and Livingston provide a lucid presentation of the chromospheric scale height problem within the context of the gaseous models [277, p. 140–145]. They recall how initial *'hydrostatic equilibrium'* arguments could only account for a density scale height of 150 km [277, p. 141]. In order to further increase this scale height to the levels observed, it was hypothesized that the chromosphere had to be heated, either through turbulent motion, wave motion, magnetic fields, or 5-minute oscillations [277, p. 140–145]. The entire exercise demonstrated that the spatial extent of the chromosphere represented a significant problem for the gaseous models. The great solar physicist Harold Zirin has placed these difficulties in perspective, *"Years ago the journals were filled with discussions of 'the height of the chromosphere'. It was clear that the apparent scale height of 1000 km far exceeded that in hydrostatic equilibrium. In modern times, a convenient solution has been found — denial. Although anyone can measure its height with a ruler and find it extending to 5000 km, most publications state that it becomes the corona at 2000 km above the surface. We cannot explain the great height or the erroneous models... While models say 2000 km, the data say 5000"* [193].

Obviously, a gas cannot support itself [62]. Hence, the spatial extent of the chromosphere constitutes one of the most elegant observations relative to the existence of a condensed solar photosphere. Within the context of the LMH model [35, 39], the Sun possesses a condensed surface. This surface provides a mechanism to support the chromosphere: gas pressure (see Fig. 17) — the same phenomenon responsible for the support of the Earth's atmosphere [48].

It was demonstrated in §4.1, that electron gas pressure cannot prevent a gaseous star from collapsing onto itself, being that these objects lack real surfaces. However, a liquid metallic hydrogen Sun has a real surface, at the level of the photosphere. When a gaseous atom within the solar atmosphere begins to move towards the Sun, it will eventually strike the surface. Here, it will experience a change in direction, reversing its downward vertical component and thereby placing upward pressure on the solar atmosphere, as displayed in Fig. 17. Gas pressure can simply account for the spatial extend of the chromosphere in condensed solar models [35, 39]. Moreover, under this scenario, the chromosphere might be supported by the escape of gaseous atoms from the solar interior as manifested in solar activity (see §5.1). This provides an acceptable mechanism in the condensed models, as they do not need to maintain the hydrostatic equilibrium essential to the gaseous Sun. In any event, chromospheric heat-

*The anchoring of coronal holes was first presented as the 22nd line of evidence [52], while the rigid rotation of the corona was once treated as the 33rd [62]. These two proofs, being closely related to one another, have now been combined.

†See the wonderful Fig. 106 in [1, p. 310, V. I] relaying the corona during the eclipse of July 8, 1842

‡This proof was first presented as the 34th line of evidence [62].

ing, from turbulent motion, wave motion, magnetic fields, or 5-minute oscillations [277], is not required to support the great spatial extent of the chromosphere in the LMH model.

4.8 Chromospheric Shape #24

Secchi had observed that the diameter of the observable Sun varied with filter selection (blue or red) during a solar eclipse [1, p. 320, V. I]. Currently, it is well established that the dimensions of the chromosphere are perceived as vastly different, whether it is studied in Hα, or using the HeII line at 30.4 nm [243, Fig. 1]. The chromosphere also appears to be prolate [243]. This prolateness has been estimated as $\Delta D/D = 5.5\times10^{-3}$ in HeII and 1.2×10^{-3} in Hα — more extended in polar regions than near the equator [243]. The shape of this layer has been demonstrated to be extremely stable, with no significant variation over a two year period [243].*

The prolate nature of the chromosphere and the extended structure which the Sun manifests above the polar axis cannot be easily explained by the gaseous models. A gaseous Sun should be a uniform object existing under equilibrium conditions, with no means of generating preferential growth in one dimension versus another. When the Sun is quiet, the greater extent of the chromosphere above the poles is associated with the presence of large anchored coronal holes in this region §(4.6). Coronal holes, in turn, manifest the presence of fast solar winds (see §5.8). A link to the fast solar winds is made in the gaseous Sun [243], despite the recognition that the origins of these winds (§5.8), and of the coronal holes with which they are associated (§4.6), remains an area of concern within these models [48, 52].

Even the oblate nature of the solar body had provided complications for the gaseous Sun (§4.4). This oblateness could be explained solely on internal cohesive forces and rotational motion in the LMH model (§4.4). But, the prolate nature of the chromosphere reflects something more complex.

According to the LMH model, fast solar winds (§5.8) are produced when intercalate atoms (see §5.1 Fig. 19) are actively being expelled from the lattice of the solar body [48, 52]. During this processes, some hydrogen is ejected, but unlike the other elements, it is often recaptured to help maintain the solar mass. In this respect, the solar chromosphere has been advanced as a site of hydrogen recondensation in the solar atmosphere (see §5.4, §5.6 and [59,61]). It appears prolate because, at the poles, more hydrogen is being expelled. Thus, more is recaptured over a greater spatial area. In analogous fashion, the corona has been designated as a site of electron recapture within the Sun [60]. With increasing distance from the solar surface, coronal atoms are increasingly stripped of their electrons. This is an electron affinity problem, wherein metallic hydrogen in the solar atmosphere scavenges for electrons and strips them from adjacent atoms [60]. Therefore, the chromosphere [59] and corona [60] act in concert to recapture protons and electrons, bringing them back onto the solar surface.

In §3.4, it was proposed [59] that the Hα emission is the direct result of the recondensation of atomic hydrogen, delivered by molecular hydrogen, onto larger condensed hydrogen structures, CHS, within the chromosphere. HeII emission results from the recondensation of atomic hydrogen, delivered by the helium hydride molecular cation [61], onto these structures (see §3.6).

In the lower chromosphere, neutral molecular hydrogen exists and can deliver atomic hydrogen with ease, resulting in Hα emission. However, with increasing height, it becomes more scarce, as the corona captures electrons. Once deprived of its sole electron, hydrogen cannot emit.

In contrast, with increased elevation, the helium hydride cation can become more abundant, as atomic helium can now harvest lone protons. Of course, neutral helium hydride in the ground state is not stable [256, 257]. Helium must first capture a lone proton (or first lose an electron to become He^+ and capture neutral hydrogen) to form the stable molecule. This readily occurs with increased height. Thus, HeII emissions are seen at the greatest chromospheric elevations. Since the helium hydride cation produced at these elevations can migrate towards the solar surface, one is able to observed HeII lines all the way down to the level of the photosphere.

Such an elegant account, exploiting chemical principles to understand line emission, cannot be framed by the gaseous models relative to the prolate nature for the chromosphere. This includes the possible causes for the differential spatial extent of Hα versus HeII lines (see Fig. 1 in [243]).

5 Dynamic Lines of Evidence

The dynamic lines of evidence involve time or orientation related changes in solar structure, emission, flow, or magnetic field. Along with many of the structural (§4) and helioseismic (§6) lines of evidence, they are amongst the simplest to visualize.

5.1 Surface Activity #25

The surface of the Sun is characterized by extensive activity.† The solar surface is often viewed as *'boiling'*, or as a *'boiling gas'*. But, gases and a gaseous Sun are unable to *'boil'*. Gases are the result of such actions. Only liquids can boil, while

*To fully understand this proof, it is necessary to simultaneously consider the origins of surface activity (§5.1), coronal holes (§4.6), solar winds (§5.8), Hα emission (§3.4) and HeII emission (§3.6). If the reader believes it difficult to follow, he/she may wish to move to other lines of evidence and return to this section once a more complete picture has been gained. This proof is listed as a structural proof (§3), even though it results from dynamic (§5) and spectroscopic (§3) processes, because it is expressed as the steady state appearance of the chromosphere when the Sun is quiet. In 1997, the sunspot number was near minimum and the data presented in [243] was acquired at that time.

†This proof was first presented as the ninth line of evidence [35, 36].

solids sublime.*

Since gases cannot boil, in order to explain activity on the solar surface, the gaseous models must have recourse to magnetic fields and flux tubes. In the case of sunspots (§2.3.3 [4, 40, 45]), faculae (§2.3.5 [45]), and magnetic bright points (§2.3.5), these fields are located within the solar body. In the case of the chromosphere (§5.6), flares (§2.3.8), and coronal mass ejections (§2.3.8), they arise from the corona. The arguments are fallacious, as magnetic fields themselves depend on structure for formation. Unable to account for their own existence (see §5.3), they cannot be responsible for creating such features within a gaseous medium.

The only prominent active features of the Sun, whose formation appears not to be inherently tied to magnetic fields, are granules (§2.3.4 [40, 45]). These are thought to be generated by subsurface heat which is being transported to the upper visible layers [40, 118–122]. A change in *'gas density'* is required within the photospheric vacuum.

In actuality, those who model granules in the laboratory (see [40] for a detailed review) understand that they are best represented as the products of Bénard convection [314–318], a process dominated by surface tension, not buoyancy [118, p. 116]. The gaseous models, unable to provide for a real surface on the Sun, must reject Bénard convection. The problem is further complicated with the realization that granules obey the 2D laws of structure (see §2.3.4) and that explosive phenomena, associated with *'dark dot'* formation, can be explained solely on the basis of structural considerations [126] (see §2.3.4). To add to the suspension of disbelief, proponents of the gaseous models maintain that the photosphere exists at the density of an ultra-low pressure vacuum ($\sim 10^{-7}$ g/cm^3 [148]). With respect to surface activity, all efforts by the gaseous models to understand the observed phenomena can be seen to collapse, when faced with the simple challenge that their solar surface is only an *'illusion'* [4]. Scientists are confronted with the intellectual denial of objective reality.

The LMH model [35, 36] can account for solar activity, since it allows for structure and takes advantage of the consequences. Granular convection can be explained with ease, as a LMH Sun possesses a true surface and the associated tension required for Bénard convection [314–318].

The emissive behavior of the Sun (see §2.3) strongly argues that the photosphere is comprised of a layered structure much like that found in graphite (see Fig. 2) and first proposed in metallic hydrogen [39] by Wigner and Huntington [88]. Layered materials like graphite are known to form intercalation compounds [48, 79–83] when mixed with other elements (see Fig. 19). In the case of metallic hydrogen, this implies that the non-hydrogen elements occupy interlayer lattice points [48], while the hexagonal hydrogen framework remains intact. It is the science of intercalation compounds which is most closely linked to the understanding of solar activity [48].

Within graphite, the diffusion of elements across hexagonal planes is hindered (see [48] for references), while diffusion within an intercalate layer is facilitated. The same principles are being invoked within the layered metallic hydrogen layers thought to exist in the Sun. Graphite intercalation compounds [79–83] are known to undergo exfoliation, an often violent process (see [79, p. 9] and [83, p. 406], where sudden phase transitions in the intercalation region from condensed to gaseous results in the expulsion of the intercalate atoms. In the laboratory, exfoliation can be associated with a tremendous expansion of lattice dimensions, as the gaseous expansion of the intercalate layers acts to greatly increase the separation between groups of hexagonal planes [79–83].

It is the process of exfoliation which can guide our understanding of solar activity. Exfoliation can be seen to result in the active degassing of the intercalation regions existing within the Sun. When the Sun is quiet, it is degassing primarily at the poles. This results in the fast solar winds (see §5.8) and coronal holes (see §4.6 [52]) in this region. It leads to the conclusion that the hydrogen hexagonal planes in the polar convection zones† tend to be arranged in a direction which is orthogonal to the solar surface.

However, in the equatorial convection zones, the hexago-

*Descriptions of a Sun which is *'boiling'* can be found throughout the printed word. Examples occur in 1) children's books [299], 2) popular writings [300, 301], 3) university level communications [302–305], 4) scientific news articles [306, 307], or 5) scholarly publications [115, 308–313]: 1) *"The sun is a boiling mass of hot gasses"* [299, p. 21], 2) *"It shows rather clearly that the Sun is a boiling mass of energy, vastly violent and constantly changing"* [300]; *"Convection is also at work transferring energy from the radiative zone to the photosphere, with a vertical boiling motion"* [301], 3) *"The surface of the Sun shows us a pattern of boiling gas arranged in a distinctive cellular pattern known as granulation"* [302]; *"Solar plasma emitted from the Sun is a boiling off of the Sun's atmosphere"* [303]; *"It is easy to think of the sun as benign and unchanging, but in reality the sun is a dynamic ball of boiling gases that scientists are only beginning to understand"* [304]; *"Our Sun is an extremely large ball of bubbling hot gas, mostly hydrogen gas"* [305], 4) *"We don't yet have a model that explains these hills"* [Jeffrey R.] *Kuhn said, although he suspects that they are caused by the interaction of boiling gas and the sun's powerful magnetic field"* [306]; *"The researchers found that, as expected, this tumultuous region resembles a pot of boiling water: hot material rises through it, and cooler gases sink"* [307], 5) *"Under poor to fair seeing conditions, sometimes the solar limb appears boiling, this gives some idea about the degree of air turbulence"* [115, p. 54]; *"The surface of the Sun boils in an active manner as the result of the continuous production of energy inside the Sun"* [308]; *"The hot corona boiling off the surface of the Sun toward the cold void of interplanetary space constitutes the solar wind"* [309]; *"The current general idea on the global balance ... is that energy conducted down from the low corona must 'boil off' mass from the chromosphere ..."* [310]; *"Near its surface, the Sun is like a pot of boiling water, with bubbles of hot, electrified gas — actually electrons and protons in the forth state of matter known as "plasma" — circulating up from the interior, rising to the surface, and bursting out into space"* [311]; *"The sun is a churning mass of hot ionized gas with magnetic fields threading their way through every pore and core, driven by energies boiling out from the interior where the fusion of hydrogen into helium at a temperature of 15 million K liberates the nuclear energy that keeps the cauldron boiling"* [312]; *"The magnetic field guides these flows, thus influencing on the average the radial distribution in the 'boiling' layer"* [313].

†A solar layer beneath the photosphere.

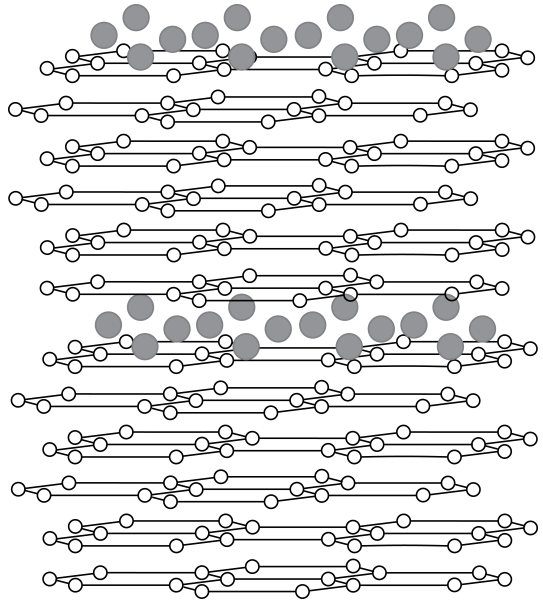

Fig. 19: Schematic representation of a proposed metallic hydrogen intercalation compound, wherein protons occupy the hexagonal lattice planes and non-hydrogen elements are located in the intercalation region. Intercalation compounds are characterized by a *'stage index'*, n, which accounts for the number of hexagonal planes between intercalate layers. In this case, n=6. This figure was previously published as Fig. 3 in [48].

nal hydrogen planes are hypothesized to be oriented parallel to the solar surface. Under the circumstances, atoms in the intercalation regions cannot freely diffuse into the solar atmosphere. They remain essentially *trapped within the Sun*, as reflected by the presence of slow solar winds above the equator. Over half the course of the eleven year solar cycle, intercalate elements slowly increase in number until, finally, the Sun becomes active (see Fig. 15) and exfoliative processes begin. The intercalate atoms begin to break and displace the hexagonal hydrogen planes, as they work their way beyond the confines of the photosphere. Coronal holes become visible at random locations throughout the Sun, indicating the reorientation of hydrogen planes in the interior. With time, the Sun degasses its equatorial region and returns to the quiet state.

In this regard, the series of images displayed in Fig. 15 are particularly telling, as they illustrate that helium levels in the lower solar atmosphere increase significantly with solar activity (examine carefully the periphery of the central image obtained in 2001 compared with images obtained in 1996 or 2005).* The Sun appears to be degassing helium, as previously concluded [48]. This further strengthens the argument that it does not, as popularly believed, possess large amounts of helium in its interior (see [47] for a detailed discussion). Rather, careful observation of the solar cycle reveals that the Sun must be comprised primarily of hydrogen, as it constantly expels other elements from its interior. The notable exception, as was seen is §3.3, relates to lithium [54].†

Relative to solar activity, the liquid metallic Sun allows for the buildup of true pressure in its interior, as intercalate elements enter the gas phase. This could account for changes in solar dimension (§4.3) and shape (§4.4, §6.3) across the cycle. It also explains the production of solar flares in accordance with ideas coined long ago by Zöllner [3, 189]. In a robust physical setting, mechanical pressure is all that is required, not energy from the corona. The same can be said of prominences, whose layered appearance (Fig. 20) highly suggests that they are the product of exfoliative forces within the Sun. Prominences reflect the separation of entire sheets of material from the Sun, exactly as found to occur when exfoliative forces act within graphite [48].

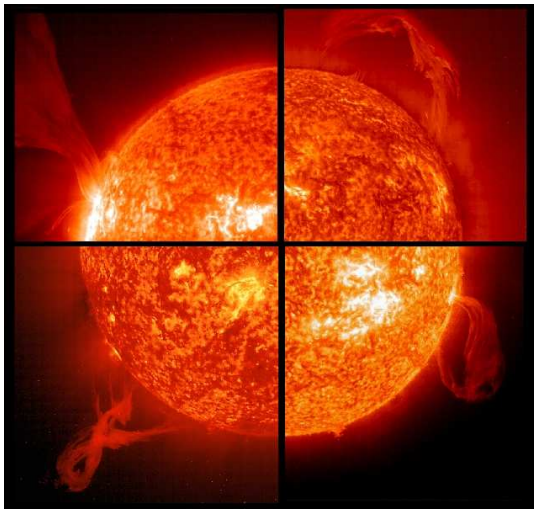

Fig. 20: An assembly of solar images obtained in the HeII line at 30.4 nm displaying the layered appearance of prominences. NASA describes this image as follows, *"A collage of prominences, which are huge clouds of relatively cool dense plasma suspended in the Sun's hot, thin corona. At times, they can erupt, escaping the Sun's atmosphere. For all four images, emission in this spectral line of EIT 304Å shows the upper chromosphere at a temperature of about 60,000 degrees K. The hottest areas appear almost white, while the darker red areas indicate cooler temperatures. Going clockwise from the upper left, the images are from: 15 May 2001; 28 March 2000; 18 January 2000, and 2 February 2001."*. Courtesy of SOHO/[EIT] consortium. SOHO is a project of international cooperation between ESA and NASA. (http://sohowww.nascom.nasa.gov/gallery/images/promquad.html — Accessed on 9/20/2013).

*Best performed using the high resolution image on the NASA SOHO website: http://sohowww.nascom.nasa.gov/gallery/images/large/304cycle.jpg.

†Deuterium and tritium, as hydrogen isotopes, should remain in the hexagonal proton planes. Like lithium, within a LMH model of the Sun, they should be retained within the solar body, with only small numbers escaping in the solar winds.

5.2 Orthogonal Flows #26

The orthogonal nature of material flow in the photosphere and corona (see Fig. 21) provides one of the simplest and most elegant lines of evidence that the Sun is comprised of condensed matter.* In 1863, Carrington established the differential rotation of the photosphere [67, 68]. His studies revealed that solar matter, at the level of the photosphere, experiences a net displacement in a direction parallel to the solar surface. Yet, solar winds (§5.8) are moving radially away from the Sun. This orthogonal flow of matter at the interface of the photosphere and the atmosphere just above it demands the presence of a physical boundary. Such a surface is unavailable in the gaseous models, but self-evident in a liquid metallic hydrogen setting.

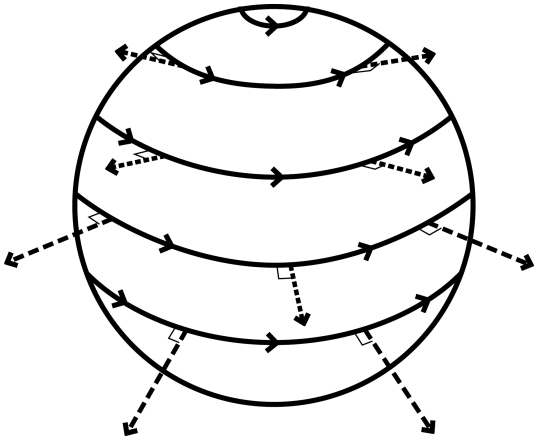

Fig. 21: Schematic representation of the orthogonal photospheric and coronal flows associated with Carrington's differential rotations [67] and the solar winds.

5.3 Solar Dynamo #27

As first noted by George Ellery Hale [107], the Sun possesses strong magnetic fields which can undergo complex windings and protrusions [12].† Magnetic fields are ubiquitous on the solar surface and within the corona. They are not manifested solely in sunspots (§2.3.3). As seen in §2.3.5, strong fields can be observed in faculae and magnetic bright points, while weak fields are present above the granules (§2.3.4) and in coronal structures (§2.3.8).

Within the context of the gaseous models, solar magnetic fields are believed to be produced by the action of a powerful solar dynamo [319, 320] generated at the base of the convection zone near the tachocline layer, well beneath the solar photosphere [12]. A dynamo represents a self-sustained amplification of magnetic fields, produced in conjunction with flow in conducting fluids. In the laboratory, they are studied using liquid metals, typically molten sodium [321–324]. Dynamo behavior must always involve the flow of conductive fluids across magnetic fields. This, in turn, *"induces electrical currents, which, under appropriate flow and magnetic field configurations, can sustain the field against dissipation"* [319].

Perhaps the greatest driving force for understanding the behavior of dynamos in the laboratory has been the presence of planetary and stellar magnetic fields [319–324]. It is not reasonable to apply these studies to a gaseous Sun.

All dynamo laboratories rely on the use of molten sodium. This substance acts as an incompressible conductive liquid metal [321–324].‡ To generate dynamo effects under experimental conditions, flow is typically induced into the metal using mechanical devices like pumps or turbines [321–324]. External induction coils are present which can provide initial magnetic fields to help either "seed" or "drive" the studies [321–324].

It is important to note that macroscopic structure is being imposed in these systems. In every case, the flow of liquid metallic sodium is being confined and directed by structure (tubes, vats, canisters) [321–324]. Insulating materials are always present, whether provided by the presence of pressurizing argon at 80 p.s.i. in a vat [321, 322] or by the inability of molten sodium to direct its own flow when propelled through pipes [323, 324]. Experimental geometries are carefully selected (see e.g. [323, Fig. 1]), including the location of induction coils [321, 322]. Mechanical devices are providing energy to drive these systems and external static magnetic fields supplement the sampling.§

In this respect, Lowe and Wilkinson constructed the first working model of a geomagnetic dynamo [328]. It was composed of solid iron alloy cylinders, rotating within a casting of the same material, wherein a small amount of mercury maintained the required electrical contact [328]. In relaying this design, Lowe and Wilkinson insisted that, *"Self-exciting dynamos are very common on the surface of the Earth, but these rely on the insulation between wires to direct the induced currents into an appropriate path; they are multiply connected"* [328].

These conditions are unlike those in gaseous stars which, by their very nature, are devoid of structure, have no ability to *"direct the induced currents into an appropriate path"* [328], and are incapable of acting as insulators. The situation has been summarized as follows, *"Whereas technical dynamos consist of a number of well-separated electrically conducting parts, a cosmic dynamo operates, without any ferromagnetism, in a nearly homogeneous medium"* [324]. With these

*This proof was first presented as the tenth line of evidence [35, 36].

†This proof was first presented as the twelfth line of evidence [35].

‡Conveniently, the density of liquid metallic sodium (ρ ~0.927 g/cm^3 [325, p. 4–128]) approaches that hypothesized to exist at the tachocline layer in the gaseous models of the Sun (ρ ~0.2 g/cm^3 [326]).

§Much like in medicine, where MRI can be performed using only the Earth's magnetic field (~0.5 gauss) [327], it is impossible to perform dynamo experiments within the laboratory in the absence of an initial ambient static field magnetic field, as has been recognized (e.g. [323]).

words, astrophysical dynamos fell outside the realm of experimental science, precisely because they are thought to exist in objects, like gaseous stars, unable to impart a physical architecture.

Astrophysics cannot hope that magnetic fields impart *'illusionary'* details and emissive properties to photospheric objects (e.g. sunspots and faculae), while at the same time requiring that real structure exists in a gaseous Sun. This structure must somehow enable the formation of powerful magnetic fields and the buildup of a solar dynamo. The fact remains that the generation of strong magnetic fields on Earth always requires the action of condensed matter. As they have no structure, gases are unable to generate magnetic fields on a macroscopic level. They are simply subject to their action. It is improper to confer upon gases behavior which cannot even be approached in the laboratory.

It is hard to envision that hydrogen in non-metallic form, as is currently hypothesized to exist in the gaseous stars, will be able to match the conductivity observed in a real metal (see Fig. 2 in [329]). Gases obviously cannot possess conduction bands and, therefore, lack the central element required to generate powerful magnetic fields on Earth. At the melting point, liquid sodium has a conductivity ($\sim 10^7\,\Omega^{-1}\,\mathrm{m}^{-1}$ [321–324]) which very much approaches that observed in the solid [321–324]. Near this point and in the solid state, conduction bands are responsible for the conductivity measured in sodium.* Hence, it should not be surprising that, just as the metal melts, some quantum mechanical conditions involved in forming these conduction bands remains (i.e. there remains some interatomic order). Otherwise, a substantial change in conductivity would be evident.

With all these factors in mind, it is reasonable to suggest that the structural lattice present in liquid metallic hydrogen provides a superior setting to account for dynamo action in the Sun. Metallic hydrogen should be able to support real structure. Protons would occupy the hexagonal planes (see Fig. 2) and electrons flow in the conduction bands necessary to generate magnetic fields. A LMH Sun should display a density, throughout its interior, similar to molten sodium. Conductive paths could be set up in the hexagonal hydrogen (i.e. proton) planes which can benefit from the insulating action of intercalate elements (see Fig. 19). As a direct consequence, changes in the dynamo and in the magnetic field intensity, in association with the solar cycle, can be accounted for as a byproduct of exfoliative forces (see §5.8). When the intercalate elements are expelled from the Sun, conductive shorts are created between hexagonal hydrogen planes which were once insulated from one another. This provides a mechanism to both build and destroy the solar dynamo. Furthermore, by turning to this substance as a solar building block, laboratory dynamo experiments become linked to a substance which may come to have great importance on Earth [92, 98], not only in the distant stars.

*Thermal vibrations can lower conductivity as temperatures are increased, but this effect is neglected in this case since both solid and liquid phases can exist at the melting point. Thus, any effect of thermal vibrations should be similar at this temperature in both phases.

5.4 Coronal Rain #28

Innocuous findings can lead to the greatest discoveries.† In this respect, coronal rain [330–333] will not present an exception. This subtle effect consists of *"cool and dense matter"* which is *"ubiquitous"* within the solar atmosphere and which is constantly falling towards the solar surface [330–333]. It is said to be composed of a *"a myriad of small blobs, with sizes that are, on average 300 km in width and 700 km in length"* [333]. When these aggregate, they produce *showers* [333]. Coronal rain has been associated with coronal loops and attempts have been made to link its existence to loop substructure [334].

As coronal rain falls towards the surface, its rate of descent does not match that expected from gravity considerations alone [333]. From the standpoint of the gaseous solar models, it appears that coronal rains and showers are retarded by the effects of gas pressure in the solar atmosphere [333]. These models rely on cycles of heating and condensation to explain coronal rain [332, 333]. But these arguments are not consistent with the belief that the lower chromosphere has a density of only $\sim 10^{-12}$ g/cm^3 [115, p. 32] and that gas pressure cannot exist (§4.1) in these models. How can condensation take place within a hot corona (see §3.7) while maintaining a gaseous state, which even at photospheric densities, would only be $\sim 10^{-7}$ g/cm^3 [148]? How can a vacuum retard the rate of descent of these particles? With respect to the existence of coronal rain, the gaseous models of the Sun simply lack the necessary flexibility to provide a reasonable account of this phenomenon.

Alternatively, the LMH model [35, 39], has advanced that condensed matter populates the outer solar atmosphere (see §2.3.6, §2.3.7, §2.3.8, §3.4, §3.5, §3.6, §3.8, §4.6, §4.7, §4.8, §5.5, §5.6, §5.7, and §6.6). Cool/dense coronal and chromospheric layers consequently stand as pillars of this model [56–60]. In this regard, the presence of coronal rain can be more readily explained if one permits true condensation to occur within the solar atmosphere.

As highlighted in §2.3.7 and §2.3.8, the K-corona should be viewed as a region containing diffuse metallic hydrogen [57, 60]. However, given the lack of pressure which exists in the K-corona, this metallic hydrogen cannot regenerate itself. Rather, coronal metallic hydrogen has entered the solar atmosphere after being expelled from the solar body during active periods (see §2.3.8, §5.5, §6.6 and [57, 58, 60]).

Though coronal LMH would be unable to self-regenerate, it should be able to provide a surface upon which other materials could condense. This appears to be what is happening with coronal rain.

†This proof was first presented as the 23rd line of evidence [53].

In this regard, it is important to note that coronal rain is usually visualized in Hα and CaII [334]. These emission lines are chromospheric in nature (see §3.4 and §3.5). Their use in detecting coronal rain strongly suggests that this material, unlike the coronal loops (§5.5) with which it is often associated [334], is actually condensing chromospheric material.*

Thus, much like water vapor on Earth condenses in the morning on the grass, hydrogen, in non-metallic form, appears to generate a dense condensate onto the coronal metallic hydrogen framework. This could explain why coronal rain can been seen flowing down coronal loops [334]. As the two substances are distinct, the hydrogen condensate slowly drifts back down to rejoin the solar surface. Since coronal rain remains attracted to the metallic hydrogen surfaces of the corona, it is unable to simply respond to the forces of gravity and its descent appears to be retarded.

Consequently, the analysis of coronal rain and its behavior appears to provide wonderful examples of the interplay between structure and function within the solar atmosphere. It strongly suggests that two distinct forms of condensed hydrogen are present in this region: 1) dense molecular hydrogen in the chromosphere [92] and 2) metallic hydrogen in the corona. Coronal rain is assisting in the harvest of hydrogen atoms from the corona. In unison, the metallic hydrogen framework, upon which it is condensing, acts to scavenge electrons from non-hydrogen atoms [56–60], which it could channel either to the solar body, or directly to coronal rain. In this manner, the corona functions to help preserve both the mass and charge balance of the Sun.

5.5 Coronal Loops #29

Coronal loops can be readily observed, both in the continuum [178–180] (see §2.3.8) and using distinct atomic emission lines (see §3.5 and §3.6), as shown in see Fig. 22. They represent *"inhomogeneous structures"*, which appear to be attached to the solar surface and which can extend well into the outer atmosphere [335, p. 83–84]. They can be relatively small (1 Mm in length and 200 km thick) or have great physical extent (several million meters to *"a substantial fraction of the solar radius"* with diameters of 1.5 Mm) [336]. While loops do not seem to possess substructure at the resolutions currently available [336], they may display such features on scales of about 15 km [336], a value well beyond current resolutions. Based on the analysis of coronal rain, it has been suggested that coronal loops have substructures smaller than 300 km [334].

As discussed in §5.4, coronal loops are associated with the presence of coronal rain. In this regard, the former may well represent a metallic hydrogen framework within the solar atmosphere unto which chromospheric matter, like coronal rain, can condense. This would appear to be confirmed

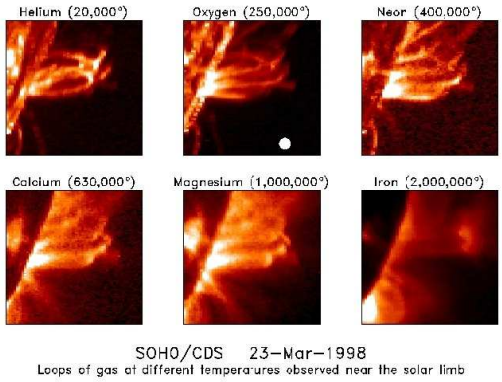

Fig. 22: Coronal loops visualized in helium, oxygen, neon, calcium, magnesium, or iron. Temperatures associated with each image have been inferred from the gaseous solar models. They correspond to 20,000 K, 250,000 K, 400,000 K, 630,000 K, 1,000,000 K, and 2,000,000 K, respectively. NASA describes this image as follows, *"CDS can produce images of the Sun at many wavelengths. In addition to hydrogen, the Sun's atmosphere contains atoms of common elements like helium, oxygen and magnesium. In the high temperature conditions of the Sun's atmosphere, these atoms emit light at different wavelengths depending on the temperature of the gas containing them. Therefore by tuning into different wavelengths we can make images of material which is at different temperatures. This capability is illustrated in the picture above, where CDS has taken images of magnetic loops of material which extend high into the Sun's atmosphere. These loops have been rendered more easily visible by observing them when they occur near the limb of the Sun, and hence they are highlighted against the dark background of space. The elements and their characteristic temperatures are indicated on the individual images. One of the surprises that the new SOHO/CDS data have produced is to show that loops at different temperatures can co-exist in the same regions of the Sun's atmosphere. The white disk plotted on the oxygen image shows the Earth to the same scale."* Courtesy of SOHO/[CDS] consortium. SOHO is a project of international cooperation between ESA and NASA. (http://sohowww.nascom.nasa.gov/gallery/SolarCorona/cds015.html — Accessed on 9/29/2013).

in Fig. 22, as both chromospheric lines (see §3.4, §3.5, §3.6) and coronal lines (see §3.8) can be detected within coronal loops.

Coronal loops hold an interesting line of evidence for condensed matter. It has been observed that *"the hydrostatic scale height. . . has always the same vertical extent, regardless of how much the loop is inclined, similar to the water level in communicating water tubes with different slopes"* [335, p. 84] (see Fig. 23).

The vertical height to which some coronal loops appear filled with matter does not change depending on inclination. The loop is containing matter which behaves as a liquid. Conversely, if the loop was merely plasma, the effects of vertical extent on loop appearance would be difficult to justify.

In this regard, it may well be that the manner in which

*Chromospheric matter is likely to be comprised of condensed matter where molecular interactions between hydrogen atoms persist [92].

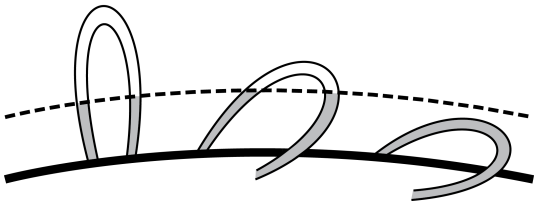

Fig. 23: Schematic representation of the vertical extent of scale height (dashed line) in coronal loops. Material fills the loop up to the scale height. If the loop is significantly inclined from the vertical axis, then it can be somewhat evenly filled with matter. The analogy can be made with water filling a tube which is more or less inclined [335, p. 84].

coronal loops appear to *'fill'* with height might represent a build up of condensed hydrogen onto these structures. As the loops assume an increasingly vertical position, material of a chromospheric nature should slowly settle towards the base of these structures, as it makes its descent down to re-enter the solar interior (see §3.4, §3.5, §3.6). Gaseous solar models are unable to rival this explanation.

5.6 Chromospheric Condensation #30

As discussed briefly in §3.4, the chromosphere is filled with spicules [337] which seem to extend as disoriented hair beyond the surface of the Sun.* As demonstrated in Fig. 24, spicules can be observed in Hα. They can also be seen in other chromospheric emission lines, including those from calcium and helium (see §3.5, §3.6 and [150, p. 8]).

The gaseous models of the Sun have no simple means to account for the formation of these structures.† Proponents of these models have expressed that two classes of spicules exist. Type II spicules are short-lived (10-150s), thin (<200 km), and said to fade [338]. Type I spicules have a 3–7 minute lifetime and move up and down [338]. It has been stated that Type II spicules might be responsible for heating the corona [338], but this claim, along with the very existence of Type II spicules, has been challenged [339]. Nonetheless, despite the densities brought forth, spicules are still believed to be propelling matter into the corona.

Counter to these ideas, the metallic hydrogen model holds that spicules are the product of condensation reactions (see §3.4, §3.5, §3.6 and [59, 61]). They enable hydrogen atoms, gathered in the solar atmosphere, to rejoin the solar body. The greatest clues for such a scenario come from the analysis of spicular velocities which appear to be essentially independent of gravitational forces [209–215].‡

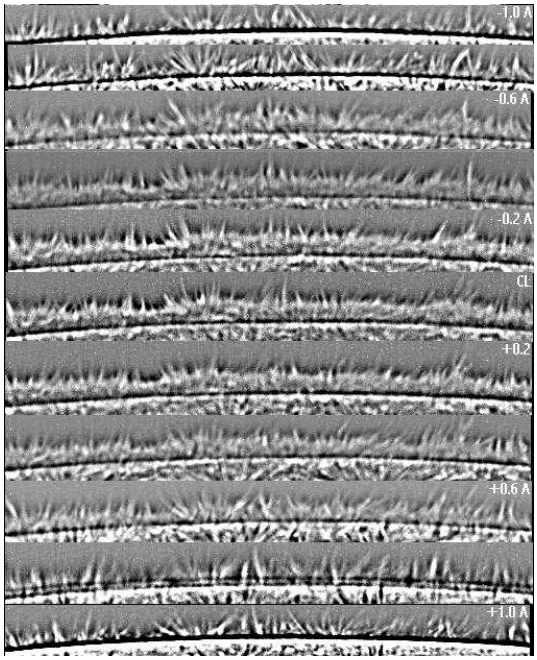

Fig. 24: A series of images displaying spicules in Hα on the solar limb. These images are displayed through the courtesy of the Big Bear Solar Observatory which have described the series as follows, *Limb Spicules: The Figure shows the limb of the Sun at different wavelengths within the H-alpha spectral line (from 0.1 nm bluewards to 0.1 nm redwards of the line center). Some of the spicules (jets) extend above height of 7000 km. The images have been processed with a high pass filter.*" http://www.bbso. njit. edu/images.html — Accessed on 9/30/2013.

Spicules seem to move up with nearly uniform speeds [206, p. 61]. These speeds can actually increase with elevation [150, p. 45–60]. Spicules can rise in jerky fashion or stop quite suddenly [150, p. 45–60]. They can *"expand laterally or split into two or more strands after being ejected"* [337].

All of this behavior, and the ability to document it, suggests that spicules are not devoid of density against an even sparser background. Rather, they seem to be the product of condensation. It is almost as if much of the material in the chromosphere exists in a state of critical opalescence, that strange state wherein matter is not quite liquid and not fully gaseous [35].§ Just a slight disturbance can cause the entire substance to rapidly condense. Such a process would be essentially independent of direction (vertical or horizontal), but

*This proof was first presented as the seventh line of evidence [35, 56, 59, 61].

†Spicules extend well into the lower corona where densities, according to the gaseous models, could be no greater than $\sim 10^{-15}$ g/cm^3, i.e. the density of the upper chromosphere [148]. The associated densities are $\sim 10^{-12}$ of the Earth's atmospheric density at sea level ($\sim$1.2x10^{-3} g/cm^3 [149]).

‡Some authors have attempted, although not very convincingly, to establish a relationship between spicular velocities and gravitational forces (e.g. [337]).

§The author has previously described the situation as follows, *"Critical opalescence occurs when a material is placed at the critical point, that combination of temperature, pressure, magnetic field, and gravity wherein the gas/liquid interface disappears. At the critical point, a transparent liquid becomes cloudy due to light scattering, hence the term critical opalescence. The gas is regaining order as it prepares to re-enter the condensed phase"* [35].

would be guided by local fluctuations in material concentrations. This would explain the erratic behavior and orientation of spicules.

The formation of spicular material suggests processes that are being observed near the critical point of a dense form of hydrogen [92] in the chromosphere. In moving from the corona to the photosphere, the effect of gravity becomes more important and, though temperatures might not be changing much (see §2.3.7), material in the chromosphere could be falling sufficiently below the critical point to allow for rapid condensation [35].*

Whether or not critical phenomena are being expressed in the chromosphere [35], it remains relatively certain that spicules themselves represent sites of condensation in the solar atmosphere, as manifested both by their dynamic behavior and by the emission lines with which they are associated (§3.4, §3.5, §3.6 and [59, 61]). It is highly likely that spicules are not propelling matter into the corona, but rather, that they are enabling hydrogen, present in the solar atmosphere, to reassume a condensed state and return to the solar body. In this case, they act to harvest hydrogen and return it to the photospheric intergranular lanes [59], as illustrated above in Fig. 14.

As with coronal rain, the chromospheric matter which makes up spicules should be comprised of dense hydrogen which is non-metallic, as it retains some hydrogen-hydrogen molecular interactions within its lattice [92]. This dense form of hydrogen, upon entering the pressurized environment of the solar interior, could then be transformed back to the metallic state [59].

5.7 Splashdown Events #31

Following violent flares, matter can be seen falling, in large fragments, back onto the solar surface.† The phenomenon resembles a huge mass of liquid projected into the air and then crashing back to the ground. A particularly impressive event was witnessed on June 7, 2011 [340, 341]. Solar material was ejected, as a great, almost volcanic appearing event, occurred on the photosphere. Solar matter was projected far into the corona, reaching heights well in excess of 500,000 km. Upon reaching a certain impressive altitude, the ejected photospheric matter was seen to fall back onto the solar body. Striking the surface, the descending material produced strong brightening at the impact points.

These events elegantly support the contention that flares and CMEs are driven by the buildup of pressure within the solar interior, not by transferring energy from the corona [189]. Most importantly, following the ejection of material from a flare, the return of mass towards the solar surface can be distinctly visualized. The associated impact points provide clear evidence that the ejected material and the surface upon which it splashes are comprised of condensed matter.

5.8 Solar Winds and the Solar Cycle #32

Solar winds have presented astronomy with a wealth of information, especially when addressing variations in helium abundances [342–351].‡ Two kinds of solar winds can be monitored. They are known as slow (<400 km/s) and fast (400–800 km/s) winds [349]. They differ only slightly in their particles fluxes (2.7×10^8 cm^{-2} s^{-1} versus 1.9×10^8 cm^{-2} s^{-1}, respectively), though they can have significant variations in their proton densities (8.3 cm^{-3} versus 2.5 cm^{-3}, respectively) [349]. Fast solar winds are typically associated with coronal holes [52, 349].

For the gaseous solar models, the origin of solar winds depends on the presence of a hot corona, which thermally expands as gravitational forces decrease with distance [352]. The body of the Sun is not involved, as a gaseous Sun must remain in perfect hydrostatic equilibrium, i.e. the forces of gravity must be exactly balanced with electron gas and radiation pressure [13, p. 6–7].

In bringing forth a solution for the origin of solar winds, Parker [352] would carefully consider earlier findings [353, 354]. Biermann had studied the orientation of comet tails and concluded that coronal particles were flowing away from the solar body [353]. At the same time, Unsöld and Chapman deduced that the Sun was expelling charged particles responsible for geomagnetic storms and computed the associated densities [354]. Parker would make the logical link between these events, but required for his solution that the space occupied by coronal matter expanded as it moved away from the Sun [352]. In order to permit this expansion, he postulated that the corona must exist at millions of degrees [352]. He believed that the outer corona could remain very hot, since Chapman had calculated, a few years before [355], that ionized gases could possess tremendous conductivities. Therefore, heat could be channeled from the lower corona to the outer solar atmosphere, to drive the solar winds.

As a result, the gaseous models have required the impossible from the corona. The latter must be heated to temperatures well beyond those of the solar core (see §3.8) using processes based on magnetic fields [148, p. 239–251]. Then, it must transfer this energy in two directions. First, the corona must be able to drive all violent activity on the solar surface [12], like flares and coronal mass ejections (see §5.1 and [179]). Second, it must allow energy, through its elevated conductivity [355], to reach the outermost layers of the solar atmosphere. In this manner, the corona itself can provide the thermal energy required to drive the solar winds [352].

But, if energy can dissipate into the outer corona through elevated conductivity, how can it be available to drive surface activity? How does the directionally opposite flow of heat in a

*There could be substantial opposition to the idea that critical phenomena are being observed in the chromosphere. However, spicule formation seems to reflect the scale length effects which characterize these processes.

†This proof was first presented as the 24th line of evidence [53].

‡This proof was first presented in [47, 48, 52].

conductive material, like the corona, not constitute a violation of the Second Law of Thermodynamics?* Furthermore, why require that heat be transferred into the corona from the solar interior prior to its application elsewhere in the Sun? Why not simply let the solar body do the work?

In any event, to maintain the requirements of hydrostatic equilibrium [13, p. 6–7], the Sun must let its ultra-low density vacuum-like corona maintain every unexplained process. It does so by transferring energy from the solar interior using magnetic fields, even though gases are unable to generate such phenomena §5.3.

The requirements that the corona is hot also introduces the problem of the cool K-coronal spectrum (see §2.3.7), which must, in turn, be explained with relativistic electrons. How could relativistic electrons survive in a conductive medium? Resorting to this proposal hampers the search for the underlying causes of the solar cycle.

Conversely, Christophe Robitaille has theorized that the Sun is expelling non-hydrogen elements synthesized within its interior (private communication and [48]).† In the LMH model, the Sun possesses a true graphite-like layered lattice (see Fig. 2) over much of its volume, except perhaps, in the core.‡ It is known in graphite, that layered lattices can accommodate the intercalation of atoms [18], as has been illustrated in Fig. 19. In this case, protons occupy the hexagonal planes, electrons are flowing in conduction bands, and non-hydrogen atoms are found in the intercalation regions. These atoms can freely diffuse in the intercalation zones, but would experience restricted diffusion across hexagonal hydrogen planes (see Fig. 19). Such simple considerations, within the context of intercalate structures, can readily account for the solar winds [47, 48, 52].

In this model, the tremendous pressures within the solar interior provide the driving forces for the solar wind. Non-hydrogen atoms in intercalation regions are being expelled from the solar body by simple mechanical action, in accordance with known exfoliative processes in graphite [48]. For instance, an atom traveling at 800 km/s could leave the center of the Sun and escape at the surface in only fifteen minutes [52].§

During quiet solar periods, the known presence of fast solar winds over coronal holes [52, 349] could be readily explained. It requires that the intraplanar axis (A in Fig. 2) of metallic hydrogen, in the polar convection zone, be positioned orthogonally to the solar surface [52]. This would enable the rapid ejection of intercalate atoms from the solar interior at the poles when the Sun is quiet.¶ In the convection zone below the solar equator, the intraplanar axis (A in Fig. 2) would be rotated by 90°, becoming parallel to the solar surface. This would act to restrict the degassing of intercalate atoms, resulting in slow solar winds above the equator.

A clearer understanding of solar winds provides new insight into helium abundances [47]. It has been argued that current estimates of solar helium levels are largely overestimated [47]. Evidence suggests that, during active periods, the Sun is expelling helium from its equatorial region, not retaining it (see Fig. 15) [47].

Helium levels in the solar wind can vary substantially with activity. When the Sun is quiet, the average He/H ratio in the slow solar wind is much less than 2%, often approaching <0.5 % (see Fig. 1 in [348]). However, when the Sun is active, the ratio approaches 4.5% [348]. Relative helium abundances can rise substantially with solar activity, like flares [347], and the He/H ratio increases dramatically during geomagnetic storms [343]. Extremely low He/H ratio values of 0.01, rising to 0.08, with an average of 0.037 have been reported, when the Sun was quiet [343]. He/H ratios can vary greatly, especially in slow solar winds [343, 346]. Therefore, astronomers have assumed that solar winds cannot be used to assay this element [347]. However, it is more likely that what is being observed has not been correctly interpreted.

Extremely low He/H ratios challenge the premise that the Sun has an elevated helium abundance [47, 241, 242], sending shock waves throughout cosmology (see [47] for more detail). As helium can be essentially absent from the solar wind, astronomers, rather than infer that the Sun has a low helium abundance, assume that the elements must not be properly sampled. Helium must be gravitationally settling in the Sun (see [48] for a detailed discussion) or is being destroyed on the way to the detectors by processes occurring in the corona [347, p. 298].

The fast solar wind is thought to represent a less biased appraisal of elemental abundances [347, p.295], precisely because helium is being ejected from the Sun and subsequently appears abundant. Aellig et al. report that the fast solar wind has a helium abundance of 4–5% throughout the course of their five year observation (see Fig. 2 in [348]).

These results can be readily explained when considering that the Sun is condensed matter. When the Sun is quiet, it is degassing its intercalation regions, primarily from the poles. Large amounts of helium can accordingly populate the fast solar wind. When solar activity is initiated, the Sun begins to degas its equatorial regions. Much of this helium then travels along with slow solar winds to our detectors, and those concentrations are likewise elevated. However, when the Sun is

*It is already difficult to accept that a low density vacuum could transfer its energy to the solar surface. This scheme becomes even more strained when coronal energy is permitted to flow freely, using conductive paths, away from the Sun. The only solution implies a violation of the First Law of Thermodynamics, i.e. energy is being created in the middle of the corona.

†Lithium provides one notable exception, as seen in §3.3 and [54].

‡A body center cubic structure, as proposed in computational studies of dense plasmas by Setsuo Ichimaru [97], would be appropriate for the solar core (see §6.5).

§This compares to thousands, perhaps millions, of years for a photon to leave the core of the gaseous Sun (see §2.3.1 and [42]).

¶Coronal holes persist above the poles during periods of reduced solar activity (see §4.6).

quiet, virtually no helium reaches our detectors in the slow solar winds, as this element is now trapped in the equatorial intercalation regions. This scenario provides strong motivation for concluding that the Sun is actively degassing helium and that the true internal abundances of this element must be much lower than currently estimated [47, 241, 242].*

Not only can the LMH model account for the production of solar winds, but it advances an underlying cause of the solar cycle: degassing of the solar body [48, 52]. When the Sun is quiet, fast solar winds are able to degas the convection zones below the poles. This helps to explain why sunspots are never seen at these latitudes. However, during this period, the equatorial regions are experiencing restricted degassing. This is due to the parallel orientation of the hexagonal hydrogen planes in layered metallic hydrogen lattice, with respect to the solar surface. Such an orientation prevails in the underlying convection zone when the Sun is quiet. Solar activity is initiated when active degassing of the equatorial planes begins. This occurs in association with a rotation or partial breakdown of the hydrogen planes, as was seen when discussing sunspots (§2.3.3). This is the reason why coronal holes can appear anywhere on the solar surface when the Sun is active, as discussed in §4.6. When accounting for solar winds, coronal holes, and solar activity, the LMH model far surpasses in insight anything offered by the gaseous models.

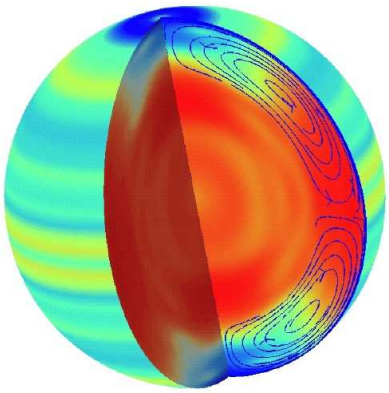

Fig. 25: Variations in sound speed within the Sun. Red regions are hotter than the standard solar models, while blue regions are cooler. This image has been provided courtesy of SOHO/[Michelson Doppler Imager] consortium. SOHO is a project of international cooperation between ESA and NASA. (http://sohowww.nascom.nasa.gov/gallery/images/mdi025.html — Accessed on 10/1/2013).

6 Helioseismic Lines of Evidence

Seismology remains a science of the condensed state. Even so, proponents of the gaseous models adhere to the belief that helioseismology can claim otherwise. In this section, a group of six helioseismic conclusions will be briefly examined. Each provides compelling evidence that the Sun is comprised of condensed matter. It might be argued that other helioseismic lines of evidence could be extracted. Only six have been selected for their scientific impact.

6.1 Solar Body Oscillations #33

The Sun acts as a resonant cavity.† It sustains oscillations, as sound waves travel (see Fig. 25), within its interior [356–360]. The most prevalent solar oscillation has a period of 5 minutes, but many more modes exist [356–360]. Thus, the solar surface is reflecting internal audio waves and this causes the entire solar body to *'ring'*, as it succumbs to seismic activity.

Though scientists currently utilize helioseismology to justify the gaseous models [356–360], the conclusions would be better suited to a condensed Sun. It is not reasonable that a photosphere, with a density of only $\sim 10^{-7}$ g/cm^3 [148], can act as a resonant cavity. Within the gaseous models, the Sun has no distinct surface, hence it cannot provide a physical boundary to sustain solar oscillations.

Fig. 25 displays slight differences in sound speed with the standard gaseous model. A detailed analysis of such studies can be profitable. Bahcall et al. [361] have also compared theoretical results with experimental helioseismic findings for standard gaseous models. Absolutely amazing fits are obtained throughout the solar interior, but the authors fail to provide comparisons for the outer 5% of the Sun (see Figs. 12 and 13 in [361]). Yet, all observational data is being acquired precisely from this region. Therefore, any perceived experimental/theoretical agreement has little validity.

As was concluded in §3.1, the Sun presents the observer with a distinct surface in the UV and X-Ray bands. This surface is covered by low-frequency 3 mHz oscillations [362]. Evidence for a distinct surface has also been presented by gamma-ray flares (see §3.2). The Sun behaves as a resonant cavity in the audio bands, implying a true surface. But the gaseous models must maintain that the solar surface is but an *'illusion'*, to somewhat poorly account for limb darkening (see §2.3.2). Unfortunately, illusions make for poor resonant cavities. It is more logical to infer that the Sun has a distinct surface over the entire span of relevant wavelengths (audio to X-ray), as provided by condensed matter.

Despite denial that the Sun is either liquid or solid, astronomers refer to solar seismic events as *"similar to earthquakes"* [362]. Such analogies are in keeping with the known truth that seismology is a science of condensed matter. The same can be said for the Sun.

*In this regard, it should be remembered that the chromosphere and the corona are working to actively recapture hydrogen, protons, and electrons. This would act to elevate the He/H ratio detected in any solar wind. In addition, since the Sun is degassing intercalate regions and its average stage index (see Fig. 19) may be quite large, the solar body might best be viewed as composed almost entirely of hydrogen.

†This proof was first presented as the fifth line of evidence [35, 36, 42].

6.2 Mass displacement #34

On July 9, 1996 a powerful X-ray flare disrupted the solar surface, as illustrated in Fig. 26 [362, 363].* This image was obtained through Doppler methods. Consequently, material moving towards the observer appeared brighter, while matter propagating away from the detector seemed darker. Therefore, the flare itself was bright.

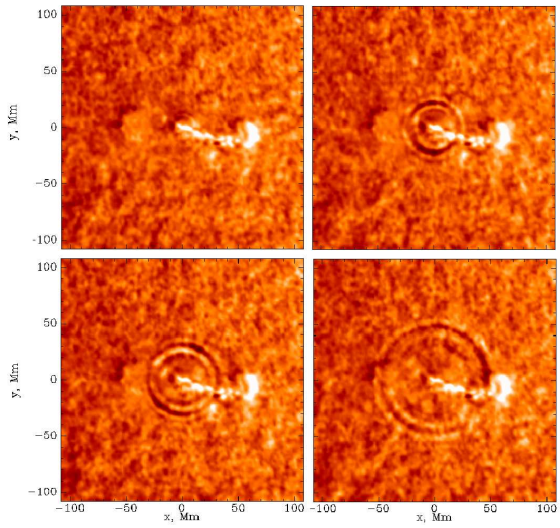

Fig. 26: Doppler image of a solar flare and the associated disturbance on the solar surface acquired by the NASA/ESA SOHO satellite [362]. Courtesy of SOHO/[Michelson Doppler Imager] consortium. SOHO is a project of international cooperation between ESA and NASA.

Kosovichev and Zharkova [362] support the notion, central to the gaseous models, that flares are being excited with coronal energy. They suggest that *"a high-energy electron beam* (is) *heating the cool chromospheric 'target'"*. Surface activity is driven, not from the interior of the Sun, but from the coronal vacuum. Nonetheless, the displacement of material observed in Fig. 25 strongly supports Zöllner's ideas regarding the nature of solar flares, as previously discussed in §5.1 and §5.7. It appears that the flare was produced when pressurized material was ejected from the solar body beyond the photospheric surface.

But, when the flare emerged, it produced enormous transverse waves on the surface of the Sun. The crest to crest distances are on the order of 10 Mm. Kosovichev and Zharkova [362] describe these transverse waves as *"resembling ripples from a pebble, thrown into a pond"* and maintain that the behavior can be explained with computations involving gas models. Still, they visualize *"ripples on a pond"*, a direct reference to behavior which can only be observed in condensed matter. Gases can sustain longitudinal, not transverse waves. Attempts to generate these waves, not only in a gas, but in an ultra-low-density vacuum, challenges scientific reason.

6.3 Higher Order Shape #35

Seismological studies have revealed that the Sun is not perfectly oblate (§4.4) but rather, is characterized by higher order quadrupolar and hexadecapolar shape terms which appear dependent on the solar cycle [364].† Higher order shape terms involve forces beyond those produced with simple rotation of a homogeneous liquid mass. They imply *internal structure within the Sun*. Hence, they stand as a sublime indication that the solar body possesses real structure beyond the core.

It would be extremely difficult to justify that fully gaseous objects could ever sustain observable internal structural effects. Yet, the higher order quadrupolar and hexadecapolar shape terms must arise from internal structure. Conversely, within the context of the LMH model, higher order shape terms would be expected. It has already been mentioned that the hexagonal hydrogen plane orientation (see Fig. 19), at the level of the convection zone, could account for coronal holes, solar winds, and the solar cycle (see §5.8). Hexagonal hydrogen planes could give rise to large layers, moving over one another, whose orientation relative to the solar surface could slowly vary from equatorial to polar regions (i.e. parallel versus orthogonal).‡ This would give rise to true underlying structure in the convection zone, as expressed in higher order shape terms.

6.4 Tachocline and Convective Zones #36

The Sun possesses a convection zone characterized by differential rotation [356–360].§ While a gas can easily be thought to undergo differential rotation, the Sun is characterized by another region: a tachocline layer separates the convection zone from the solid solar core (see §6.5).

The tachocline region acts as a shear layer within the Sun. This layer is known to be prolate in nature [360, 365–367]. The tachocline is generally thicker and shallower at the higher latitudes [360, 366]. It seems to display some temporal variability across the solar cycle [366], strongly suggesting, once again, that structural changes are taking place within the solar body (see §5.8 and §6.3).

When considering the tachocline layer, it is important to recall that shear stresses require the presence of a physical plane. For instance, the equation for shear stress, τ, states that τ=F/A, where F=force and A=Area. It is not possible

*This proof was one of the earliest [4, 29] and was presented, at one time, as the sixth line of evidence [35].

†This proof was first presented in [50], as supportive of §4.4. However, solar oblateness does not depend on the use of helioseismology for its determination (§4.4) and has been invoked by Jeans [27, 28] as providing a mechanism to generate binaries [3]. As for higher order shape, it is indicative of forces which differ from those involved in creating oblateness. Upon reconsideration, higher order shape now stands on its own as a separate line of evidence.

‡This resembles tectonic shifts on Earth. Such a parallel was drawn by Luc Robitaille (personal communication).

§This proof was first presented as the nineteenth line of evidence [50].

to have a shear stress without acting on a surface, or an organized lattice plane of atoms, as provided by condensed matter. Imaginary planes cannot experience shear forces.

Consequently, the shear nature of the tachocline, and the fact that it displays a prolate nature, provides clear evidence that the solar body is physically structured. Furthermore, it appears that this is an area of the Sun which can undergo changes with the solar cycle. These results are most gracefully explained by the LMH model.

6.5 Solar Core #37

As was suggested in §6.4, the core of the Sun undergoes solid body rotation [368].* This conclusion, has been reached by a virtual *who's who* of authority in helioseismology [368]. In the central portion of the Sun, *"...the rotation rate appears to be very little, if at all. Its value is 430 nHz"* [368].

Solid body rotation in the solar interior directly implies that the body of the Sun cannot be gaseous. This rotation requires the presence of powerful cohesive forces within the Sun. None can exist in a gaseous object.

The observation is more in line with Setsuo Ichimaru's conjecture (§2.3.1 and §5.8) that the central portion of the Sun can be considered to exist as a one-component plasma of metallic hydrogen [97, pp. 103 & 209]. Ichimaru adopted the body-centered cubic structure in his studies [97–99] and this lattice configuration would make sense at the center of the Sun.

In this respect, Ichimaru based the density of metallic hydrogen in the core on conclusions derived from gaseous models. If the photosphere of the Sun is truly condensed, then the values he adopted (56.2 g/cm^3 [98, p. 2660]) would be much too elevated. In a liquid model, the density cannot vary much throughout the solar body, remaining near 1.4 g/cm^3 (i.e. slightly lower at the photosphere and slightly higher in the core). At the center of the Sun, we are merely witnessing a change in lattice structure from a layered Type-I lattice over most of the photosphere, to a more metallic layered Type II lattice in the convection zone, and finally to a body-centered cubic lattice in the core. Intercalate atoms would be present within Type I and Type II layered lattices. If they change from the condensed to the gaseous phase, these intercalate atoms could slightly reduce the average densities of these layers.

The LMH model is more in keeping with physical observations within the Sun. It is not reasonable to advance that gases rotate as solid bodies. Condensed matter enables the formation of a solid core which can account for the observed rotations.

6.6 Atmospheric Seismology #38

Helioseismology has been extended to the outer solar atmosphere [214, 369–372].† Coronal and chromospheric studies [214, 369–372] have successfully detected seismic waves in this region of the Sun and the presence of both incompressible and compressible waves is now well-established. These are viewed as magnetohydrodynamic waves (MHD) in nature.‡

The existence of incompressible transverse waves in the solar atmosphere [214, 369–372] suggests, once again, that this region of the Sun contains condensed matter. These have been observed in spicules [214] and within the chromospheric level [372]. Their detection implies that the densities of these solar layers are well in excess of those which typify Earthly vacuums.

As a point of interest, it is known that comets can send shock waves throughout the solar corona and chromosphere. On January 29, 2013 (see [373]), a comet begins to disrupt the solar atmosphere when it is more than $1R_{\odot}$ away from the solar surface. At this location, the corona has no density ($<10^{-15}$ g/cm^3, the density of the upper chromosphere [148]), according to the gaseous models. It is unfeasible that an ultra-low-pressure vacuum could be able to respond to the entry of a comet in this manner. The ability of comets to trigger shock wave propagation throughout the solar atmosphere indicates that this is a region of elevated density. This conclusion is in keeping with the LMH model of the Sun.

7 Elemental Lines of Evidence

7.1 Nucleosynthesis #39

It has been gloriously stated that the elements were formed in the stars.§ In this, there appears to be much truth [374–388]. From its inception, stellar nucleosynthesis has always been closely linked to stellar evolution [129, 374–378].

The idea that the Sun could synthesize helium was first proposed by men such as Gamow [377, 378], Bethe [379–381], von Weisäcker [382] and Hoyle [383, 384]. The p-p reaction, wherein two protons combine to make a deuteron, while relying on positron and neutrino emission, would come to play a vital role in ^{4}He synthesis within low mass stars [374, p. 118]. For stars with a greater mass than the Sun, Bethe and von Weisäcker, in 1938 and 1939 [380–382], advanced that ^{4}He was being formed in a simple cycle involving nitrogen, carbon, and oxygen (CNO).

Early on, Hans Bethe had argued that *"no element heavier than ^{4}He can be built up in ordinary stars"* [381]. With those words, the Sun was crippled and stripped of its ability to make any element beyond helium.

Bethe had reached his conclusion based on the probability of nuclear reactions in the gas phase and at the temperatures of ordinary stellar cores [381, p. 435]. If this was true, how did the Sun come to acquire the other elements? For Bethe, the answer appeared straightforward, *"The heavier elements found in stars must therefore have existed already when the*

*This proof was first presented as the twentieth line of evidence [50].
†This proof was first presented as the 29th line of evidence [58].
‡See [372] for a brief, but well compiled, literature review.
§This proof was first presented in [44, 48].

star was formed" [381]. Extremely large and hot, first generation stars, had, soon after the Big Bang, created the heavy elements [389]. These elements merely represented contamination in the Sun, a product of objects extinguished long ago.

At the time that the CNO cycle was outlined [380–382], the discovery of metabolic cycles was creating a fury in biology. Just a few years before, in 1932, Hans Krebs (Nobel Prize, Medicine and Physiology, 1953) had discovered the urea cycle [390]. He would go on to outline the tricarboxylic acid (TCA or Krebs) cycle in 1937 [391], the discovery for which he gained international acclaim. It cannot be doubted that these great pathways in biology influenced astrophysical thought. Cycles seemed all powerful.

Biological cycles initially concealed their many lessons. It would take years to fully understand that they were highly regulated entities. Biological cycles required a complement of reactions and cofactors (small activator molecules or ions) which could either sustain the levels of intermediates or activate key enzymatic reactions. Similar regulation would be difficult to envision in the case of the CNO cycle. As a result, can this cycle truly occupy central positions in the synthesis of ^{4}He in the stars? Why confound the process by resorting to a cycle, when simple reactions between hydrogen atoms should be sufficient for all stars?

It would seem fortuitous that precisely the proper amounts of carbon, nitrogen, and oxygen has been distributed within stellar interiors, to permit these reactions to take place. If stars are truly gaseous, how do they ensure that these elements are not destroyed, or used up, by competing nuclear reactions — something which can be prevented or exploited to advantage in biology? Unlike a biological cell, with its intricate means of forming, separating, and transferring metabolites, the gaseous star cannot control the course of a single reaction. Everything must occur by chance. This complication is directly opposed to the subsistence of cycles.*

Concerning nucleosynthesis, proponents of the gaseous models require the improbable. Hobbled by theory, they must claim that first generation stars created the heavy elements. Moreover, they advance that, while mankind has successfully synthesized many elements, the Sun is unable to build anything beyond helium. First generation stars which no longer exist had done all the work [389]. These conclusions, once again, call for the suspension of disbelief. It is much more reasonable to assume that the Sun has the ability to synthesize all the naturally occurring elements, based on their presence in the solar atmosphere.

In turning his attention to dense plasmas, Ichimaru recognized that they could provide additional freedom in elemental synthesis [97–99]. These ideas have merit. In the LMH model, dense structures enable the synthesis of heavy elements which is not restricted to the solar core, but expressed in the convection zone where the intercalation regions can be found.

A metallic hydrogen framework can restrict protons to lattice points in the hexagonal plane and confine other atoms to the intercalate layer [48]. Solar pressure and lattice vibrations could act in concert to enhance the probability of nuclear reactions. Two adjacent protons, in the hexagonal hydrogen plane, could give rise to a deuterium atom, with the associated positron and neutrino emission [388]. This deuterium could then react with another, leading directly to the synthesis of ^{4}He. Alternatively, it could fuse with a proton, leading to the formation of ^{3}He. Both ^{4}He and the light helium isotope, ^{3}He, would be immediately ejected into the intercalation region [48].† Over time, the intercalation region could sustain other nuclear reactions and become the birthplace of all naturally occurring heavy isotopes. The Sun and the stars gain the ability to synthesize all of the elements [44, 48].

In this regard, it is well-known that solar flares can give tremendous ^{3}He abundance enhancements [180]. Eruptive flares have been known to produce ^{3}He/^{4}He ratios approaching 1 [186], and thousand-fold enhancements of this ratio have been observed [392]. These findings can be better understood in a solar model wherein ^{3}He is being preferably channeled into intercalation regions over ^{4}He. ^{3}He could then display an enhancement over ^{4}He when released into the solar atmosphere during activity.‡ It would be difficult to account for the finding for the gaseous models, but the result can be reasonably explained using the LMH model.§

8 Earthly Lines of Evidence

The earthly lines of evidence may be the most powerful. They are certainly the most far reaching. Climate dictates our future and the survival of humanity.

Thus, it is fitting to close this discussion with the climatic line of evidence. This acts to highlight that there is much more to studying the Sun than intellectual curiosity. As such, the *'Young Sun Problem'* and the great Maunder minimum of the middle ages are briefly discussed.¶

*Note that the author has proposed a cycle in §3.6. In this case however, the formation of triplet He has not been left to chance. It is the direct product of a systematic chemical reaction. The other reactant in the cycle, hydrogen, is present in excess.

†^{3}He could also emit a positron to make tritium, ^{3}H. Remaining in the hexagonal plane, this hydrogen isotope could then react with a single proton to make ^{4}He, which could then be expelled into the intercalate region.

‡This requires simply that the reaction of a deuterium atom with a proton is preferred over its reaction with another deuterium atom. This would be expected in a hyrogen based Sun.

§The solar neutrino problem has not been addressed in this work as a full exposition would involve too much discussion. Suffice it to state that difficulties involved in obtaining proper neutrino counts highly suggest that the Sun is sustaining other nuclear reactions beyond the simple synthesis of ^{4}He.

¶These constitute a single line of evidence as they are both related to climatic changes on Earth.

8.1 Climatic #40

8.1.1 The Young Sun Problem

The gaseous models infer that, when the Sun was young, it was much cooler than it is at present [393–395]. Once thought to be faint and dissipating much less heat onto the surface of the Earth, a gaseous Sun became increasingly warm over time. Thus, the Sun was once thought to be faint, dissipating little energy onto the Earth. Two billion years ago, the mean temperature of the Earth's surface would have been below the freezing point of water [393]. A paradox arises, since geological studies have revealed that water existed on Earth in liquid state as early as 3.8 billion years ago [393–395].

In order to resolve this problem, Carl Sagan was one of the first to advance that the answer could be found in the Earth's atmosphere [395]. If the young atmosphere was rich in CO_2, then the greenhouse effect and global warming [396] provided an explanation [393–395]. Everything appeared to be resolved [393].

Still, some remained unsatisfied with the greenhouse solution. Several stated that a young Sun was more massive and accordingly, hotter [393, p. 457]. In this scenario our Sun lost enormous amounts of material over the years through *"a vigourous, pulsation driven, solar wind"* [393, p. 457]. The young Sun could have been fifteen times more luminous than now, simply as a consequence of these changes in mass [393, p. 458].

But, it is difficult to conceive how a gaseous star, violently expelling mass despite great gravity, will cease to do so as gravitational forces decrease. Nonetheless, these basic ideas have survived, although with less dramatic changes in mass loss [397]. In this approach, the gaseous young Sun was not faint, but bright [397]. This was more in keeping with warm temperatures both on the Earth and on Mars [397]. Greenhouse effects could not simultaneously explain these findings.

In the end, the LMH model has a distinct advantage relative to the young Sun problem. Only the gaseous equations of state demand that a star like the Sun must become increasingly luminous as it evolves.* But over time, a Sun based on condensed matter, should cool from the most luminous (Class O) to the coolest star type (i.e. Class M).

Some may highlight that, if our Sun was once an O class star, there should be no water on Earth. The supposition is not valid. When the Earth was young, scientific consensus states that it was molten (see e.g. [399]). This can be easily explained if the Sun was once an O Class star, but not if it was a faint gaseous object. The Earth, like our Sun, cooled over time. The LMH model is much more in accordance with observational facts in this regard.†

8.1.2 The Maunder Minimum

A great minimum appeared in the Sunspot cycle during the middle ages. This minimum was first recognized by Spörer and Maunder [400–404]. It is known today as the *Maunder minimum* [403]. Many believe that the Maunder minimum was associated with a *'little ice age'* on Earth [403]. The conclusion is particularly timely, since the Sun may be entering another minimum in 2013, as solar activity apparently drops to a 100 year low [405].

What causes these minimae? In gaseous models, the answers will be difficult to ascertain, as these ideas have difficulty accounting for any solar activity. As for the LMH model, it is based on the tenant that solar activity must be fundamentally related to degassing of intercalate atoms. Perhaps the Maunder minimum arises because the Sun has been thoroughly degassed, either through an unknown internal mechanism or an external force.

In this regard, it may be important to recall that comets appear to send shock waves through the solar atmosphere as they come near the Sun [373]. These shock waves could be degassing our star beyond normal, hence reducing the need for future solar activity. *Shock degassing* may seem unlikely. However, comets do have periodic motions around the Sun. One or more could cyclically return to cause such effects. In this respect, the comet ISON is arriving in just a few days [406]. It will be interesting to note the shock wave it commands as it orbits the Sun.‡

8.2 Conclusion

Throughout these pages, a trial has unfolded relative to the constitution of the Sun. Prudent consideration of the question requires the objective analysis of solar data. Observations must be gathered and rigorously considered in light of known laboratory findings. Such were the lessons imparted long ago when Gustav Kirchhoff first contemplated the nature of the Sun [26].

Kirchhoff's approach has now been repeated. A wealth of information has been categorized and meticulously evaluated. Data spanning every aspect of the solar science has been included. Not a single fact was deliberately omitted or ignored. Rather, the full complement of available evidence has been weighed and described. The Sun itself was permitted to offer full testimony. In completing this exercise, a total

*The author has previously addressed Lane's law and the increased luminosity gained by the gaseous stars as they evolve [3]. With respect to stellar evolution, the LMH model will advance that stars cool as they evolve and do not increase in luminosity. The brightest stars (Classes O and A) are actually the youngest, while the faintest are the oldest (Class M). This is completely contrary to current beliefs in astronomy. Stellar evolution will be addressed in considerable in detail in an upcoming work [398].

†The mystery of the appearance of water on a planet that was once molten has not been properly addressed by anyone to the author's knowledge.

‡Shock related degassing of the Sun should be viewed as something positive. A star unable to properly degas might well exfoliate, as discussed in [48], and become a red giant or a supernova. Therefore, shock degassing may well be necessary, even if Earthly temperatures subsequently fall for rather long periods of time.

of forty lines of evidence have been addressed in seven broad categories. Each has spoken in favor of condensed matter.

Of these, the Planckian lines of evidence, as outlined in §2, will always merit the preeminent positions, since they directly reveal true lattice structure at the atomic level. The solar spectrum, limb darkening, and the directional emissivity of many structures (sunspots, granules, faculae, magnetic bright points, spicules, the K-corona, and coronal structures) highlight that metallic and non-metallic material can be found within the Sun.

The spectroscopic lines of evidence may well be the most elegant. It is not only that they provide obvious clues for a solar surface, but that they finally expose the underlying cause of line emission within the chromosphere and corona. In this regard, molecular hydrogen and the metal hydrides strongly suggest that the chromospheric flash spectrum reflects the presence of condensation reactions in the solar atmosphere. Yet, it is triplet helium which has rendered the most definitive declaration. It appears that an activated helium cycle does indeed exist in the chromosphere, harvesting hydrogen atoms and enabling them to rejoin the solar surface. In concert, the cool-LMH-containing K-corona scavenges electrons, thus helping to preserve solar neutrality. The associated light emission from highly ionized ions speaks to the power of spectroscopic observation.

The structural lines of evidence remain the simplest to understand. The many arguments concerning solar collapse, density, dimension, shape, appearance, and extent, are simultaneously straightforward and disarming.

Perhaps the most intriguing lines of evidence are dynamic manifestations of solar activity. Surface activity, the boiling action of the Sun, and the orthogonal arrangement of its photospheric/coronal flows leave no opportunity for a gaseous Sun. The existence of a solar dynamo, with its requirement for the interplay between conductors and insulators, offers no more. Coronal rain and loops, along with spicular velocities and splashdown events, require the presence of condensed matter. Slow and fast solar winds point to an object constantly striving to expel material, emphasizing the dynamic aspects of a condensed Sun.

Few sciences are more tied to condensed matter than seismology. The Sun with its oscillations, mass displacements, shape, internal layers (convection zone, tachocline, and core), and atmospheric waves, has highlighted that it belongs in the company of solids and liquids.

Elemental lines of evidence call for a complete revision of scientific thought relative to how the Sun derives its energy. First generation stars must join the company of other untenable theories, as an unchained Sun is finally permitted to synthesize all of the elements.

The sole earthly line of evidence was climatic. In ages past, the Earth was molten. The Sun must have been much more luminous than it is today, leading to the conclusion that it was born as an O-class star. Its temporal variations across the ages, might be best understood as an ever-present need to eject elements from its interior.

Finally, a conclusion must inevitably be drawn. Can a gaseous Sun truly survive, based solely on mathematical arguments, when not a single observational line of evidence lends it support? In the end, such an arsenal of observational proofs has been supplied that there can be little doubt in the answer. Formulas can never supersede observational findings. Hence, only a single verdict can be logically rendered. The Sun must be comprised of condensed matter.

The consequences are far reaching. They call for a new beginning in astronomy. Nonetheless, there is hope that a reformulation of astrophysics can bring with it a wealth of knowledge and discovery. As scientists turn their thoughts to a condensed Sun, may they renew their fervor in the pursuit and understanding of stellar observations.

Epilogue

No more appropriate closing words can be uttered than those of Cecilia Payne, she who established that we live in a hydrogen based universe [86]: "*The future of a subject is the product of its past, and the hopes of astrophysics should be implicit in what the science has already achieved. Astrophysics is a young science, however, and is still, to some extent, in a position of choosing its route; it is very much to be desired that present effort should be so directed that the chosen path may lead in a permanently productive direction. The direction in which progress lies will depend on the material available, on the development of theory, and on the trend of thought . . . The future progress of theory is a harder subject for prediction, than the future progress of observation. But one thing is certain: observation must make the way for theory, and only if it does can the science have its greatest productivity . . . There is hope that the high promise of astrophysics may be brought to fruition.*" Cecilia Payne-Gaposchkin [407, p. 199–201].

Acknowledgment

The Swedish 1-m Solar Telescope science team is recognized for Figs. 3, 7, and 9. The SST is operated on the island of La Palma by the Institute for Solar Physics of the Royal Swedish Academy of Sciences in the Spanish Observatorio del Roque de los Muchachos of the Instituto de Astrofísica de Canarias.

NASA/SDO and the AIA science team acquired the data displayed in Fig. 11. NASA/SOHO and the EIT, CDS, and MDI science teams, were responsible for Figs. 15, 20, 22, and 26. The Big Bear Solar Observatory is recognized for Fig. 24.*

Luc Robitaille is acknowledged for the preparation of all other figures.

*Agency URLs — http://www.isf.astro.su.se; http://sdo.gsfc.nasa.gov; http://sohowww.nascom.nasa.gov; http://www.bbso.njit.edu.

Not enough can be said of Dmitri Rabounski and Larissa Borissova with respect to their lifelong love of science and their immediate interest in the problem of liquid stars [408].

Dedication

This work is dedicated to those who, through their support, sacrifice, compassion, and understanding, permitted that my life be dedicated to science — my wife Patricia Anne* and our sons: Jacob,† Christophe,‡ and Luc.§

Submitted on: October 7, 2013
Accepted in revised form on: October 13, 2013
First published online on: October 13, 2013

References

1. Secchi A. Le Soleil (Vol. I/II). Gauthier-Villars, Paris, 1875/1877. (Available online: Vol. I — http://archive.org/details/lesoleil01secc; Vol. II — http://archive.org/details/lesoleil02secc).
2. Robitaille P.M. A thermodynamic history of the solar constitution — I: The journey to a gaseous Sun. *Progr. Phys.*, 2011, v. 3, 3–25.
3. Robitaille P.M. A thermodynamic history of the solar constitution — II: The theory of a gaseous Sun and Jeans' failed liquid alternative. *Progr. Phys.*, 2011, v. 3, 41–59.
4. Robitaille P.-M. On the presence of a distinct solar surface: A reply to Hervé Faye. *Progr. Phys.*, 2011, v. 3, 75–78.
5. Abbot C.G. The Sun. D. Appleton and Company. New York, 1911.
6. Maunder E.W. Are the planets inhabited? Harper & Brothers, London, 1913.
7. Fowler A. Some problems in astronomy IV. Solar and stellar photospheres. *The Observatory*, 1913, v. 36, 182–185.
8. University of Cambridge, Institute of Astronomy, Ph.D. Program, Requirements for Admission to Postgraduate Study (accessed online on 8/29/2013: http://www.ast.cam.ac.uk/admissions/phd).
9. Eddington A.S. The internal constitution of the stars. Cambridge University Press, Cambridge, U.K., 1926.
10. Fowler R.H. On dense matter. *Mon. Not. Roy. Astron. Soc.*, 1926, v. 87, 114–122.
11. Bahcall J.N. and Pinsonneault M.H. Standard solar models, with and without helium diffusion, and the solar neutrino problem. *Rev. Mod. Phys.*, 1992, v. 64, no.4, 885–926.
12. Hassan S.S. Magnetic flux tubes and activity on the Sun. In, *Lectures on Solar Physics* (H.M. Antia, A. Bhatnagar and R. Ulmschneider, Eds.), Springer, Berlin, 2003, p. 173–201.
13. Kippenhahn R. and Weigert A. Stellar structure and evolution. Springer-Verlag, Berlin, 1990.
14. Clayton D.D. Principles of stellar evolution and nucleosynthesis. MgGraw-Hill, New York, 1968.
15. Kirchhoff G. Über den Zusammenhang zwischen Emission und Absorption von Licht und. Wärme. *Monatsberichte der Akademie der Wissenschaften zu Berlin*, sessions of Dec. 1859, 1860, 783–787.
16. Kirchhoff G. Über das Verhältnis zwischen dem Emissionsvermögen und dem Absorptionsvermogen. der Körper fur Wärme und Licht. *Poggendorfs Annalen der Physik und Chemie*, 1860, v. 109, 275–301. (English translation by F. Guthrie: Kirchhoff G. On the relation between the radiating and the absorbing powers of different bodies for light and heat. *Phil. Mag.*, 1860, ser. 4, v. 20, 1–21).
17. Wien W. Über die Energieverteilung in Emissionspektrum eines schwarzen Körpers. *Ann. Phys.*, 1896, v. 58, 662–669.
18. Stefan J. Über die Beziehung zwischen der Warmestrahlung und der Temperature. *Sitzungsberichte der mathematischnatur wissenschaftlichen Classe der kaiserlichen Akademie der Wissenschaften*, Wien 1879, v. 79, 391–428.
19. Planck M. Über das Gesetz der Energieverteilung im Normalspektrum. *Annalen der Physik*, 1901, v. 4, 553–563 (English translation by ter Haar D.: Planck M. On the theory of the energy distribution law in the normal spectrum. The old quantum theory. Pergamon Press, 1967, 82–90; also Planck's December 14, 1900 lecture *Zur Theorie des Gesetzes der Energieverteilung in Normalspectrum*, which stems from this paper, can be found in either German, or English, in: Kangro H. Classic papers in physics: Planck's original papers in quantum physics. Taylor & Francis, London, 1972, 6–14 or 38–45).
20. Planck M. The theory of heat radiation. P. Blakiston's Son & Co., Philadelphia, PA, 1914.
21. Robitaille P-M. On the validity of Kirchhoff's law of thermal emission. *IEEE Trans. Plasma Sci.*, 2003, v. 31, no. 6, 1263–1267.
22. Robitaille P.-M. A critical analysis of universality and Kirchhoff's Law: A return to Stewart's Law of Thermal Emission. *Progr. Phys.*, 2008, v. 3, 30–35.
23. Robitaille P.-M. Blackbody radiation and the carbon particle. *Progr. Phys.*, 2008, v. 3, 36–55.
24. Robitaille P.-M. Kirchhoff's law of thermal emission: 150 Years. *Progr. Phys.*, 2009, v. 4, 3–13.
25. Stewart B. An account of some experiments on radiant heat, involving an extension of Prévost's theory of exchanges. *Trans. Royal Soc. Edinburgh*, 1858, v. 22, no. 1, 1–20 (also found in Harper's Scientific Memoirs, edited by J.S. Ames: The Laws of Radiation and Absorption: Memoirs of Prévost, Stewart, Kirchhoff, and Kirchhoff and Bunsen, translated and edited by D.B. Brace, American Book Company, New York, 1901, 21–50).
26. Kirchhoff G. The physical constitution of the Sun. In *Researches on the solar spectrum and the spectra of the chemical elements*. Translated by H.E. Roscoe, Macmillan and Co., Cambridge, 1862, 23–31.
27. Jeans J.H. Problems of Cosmogony and Stellar Dynamics — Being an Essay to which the Adams Prize of the University of Cambridge for the year 1917 was Adjudged. Cambridge University Press, 1919.
28. Jeans J.H. Astronomy and Cosmogony. Cambridge University Press, 1928.
29. Robitaille P.M. The collapse of the Big Bang and the gaseous Sun. *New York Times*, March 17, 2002, p. A10 (available online: http://thermalphysics.org/pdf/times.pdf).
30. Robitaille P.M. Evidence for a liquid plasma model of the Sun. *Am. Phys. Soc. Meeting — April*, 2004, S280002.
31. Robitaille P.M. The Sun as a hot liquid plasma: additional evidence. *Am. Phys. Soc. Meeting — Ohio Spring*, 2004, S50002.
32. Robitaille P.M. The photosphere as condensed matter. *Am. Phys. Soc. Meeting — Ohio Fall*, 2004, S60005.
33. Robitaille P.M. The Sun as a hot liquid plasma: more evidence. *Am. Phys. Soc. Meeting — NE Fall*, 2004, S10004.
34. Robitaille P.M. The Sun as a high energy/high density liquid metallic hydrogen plasma. *The 33rd IEEE International Conference on Plasma Science*, June 4-8, 2006, Traverse City, Michigan, p. 461, DOI:10.1109/PLASMA.2006.1707334.

*She insisted that this work be produced and that the proofs be gathered in one treatise.

†Jacob was the first to state that someday forty proofs would be published.

‡Chrisophe provided several of these lines of evidence in a paper we jointly authored based on the behavior of the solar winds and the structure of the Sun [48]. At the time, I had failed to recognize that these constituted additional proofs for condensed matter.

§Ever creative, Luc generated many of the figures in my relavent papers and has been a careful and just critic of both style and scientific presentation.

35. Robitaille P.M. The solar photosphere: Evidence for condensed matter. *Progr. Phys.*, 2006, v. 2, 17–21 (also found in slightly modified form within *Research Disclosure*, 2006, v. 501, 31–34; title #501019).
36. Robitaille P.M. A high temperature liquid plasma model of the Sun. *Progr. Phys.*, 2007, v. 1, 70–81 (also in arXiv: astro-ph/0410075v2, Oct. 4, 2004).
37. Robitaille P.M. A radically different point of view on the CMB. In, *Questions of Modern Cosmology — Galileo's Legacy*, ed. by M. D'Onofrio and C. Burigana, Springer, New York, 2009, p. 93–108.
38. Robitaille P.M. Liquid metallic hydrogen: Building block of a liquid Sun. *Am. Phys. Soc. Meeting — Ohio Spring*, 2011, D4.00005.
39. Robitaille P.M. Liquid metallic hydrogen: A Building block for the liquid Sun. *Progr. Phys.*, 2011, v. 3, 69–74.
40. Robitaille P.M. On solar granulations, limb darkening, and sunspots: Brief insights in remembrance of Father Angelo Secchi. *Progr. Phys.*, 2011, v. 3, 79–88.
41. Robitaille P.M. On the temperature of the photosphere: Energy partition in the Sun. *Progr. Phys.*, 2011, v. 3, 89–92.
42. Robitaille P.M. Stellar opacity: The Achilles heel of the gaseous Sun. *Progr. Phys.*, 2011, v. 3, 93–99.
43. Robitaille P.M. Lessons from the Sun. *Progr. Phys.*, 2011, v. 3, 100–102.
44. Robitaille P.M. Nucleosynthesis of the elements and the LMH model of the Sun. *Am. Phys. Soc. Meeting — Four Corners Annual*, 2012, D1.00026.
45. Robitaille P.M. Magnetic fields and directional spectral emissivity in sunspots and faculae: Complimentary evidence of metallic behavior on the surface of the Sun. *Progr. Phys.*, 2013, v. 1, 19–24.
46. Robitaille P.M. On the nature of the chromosphere: Condensation and line emission. *Am. Phys. Soc. Meeting — Ohio Spring*, 2013, E2.00007.
47. Robitaille P.-M. Liquid metallic hydrogen II. A critical assessment of current and primordial helium levels in the Sun. *Progr. Phys.*, 2013, v. 2, 35–47.
48. Robitaille J.C. and Robitaille P.-M. Liquid metallic hydrogen III. Intercalation and lattice exclusion versus gravitational settling and their consequences relative to internal structure, surface activity, and solar winds in the Sun. *Progr. Phys.*, 2013, v. 2, 87–97.
49. Robitaille P.-M. Commentary relative to the distribution of gamma-ray flares on the Sun: Further evidence for a distinct solar surface. *Progr. Phys.*, 2013, v. 2, L1–L2.
50. Robitaille P.-M. Commentary relative to the seismic structure of the Sun: Internal rotation, oblateness, and solar shape. *Progr. Phys.*, 2013, v. 2, L3–L4.
51. Robitaille P.-M. Commentary on the radius of the Sun: Optical illusion or manifestation of a real surface? *Progr. Phys.*, 2013, v. 2, L5–L6.
52. Robitaille P.-M. Commentary on the LMH model of the Sun: Insight relative to coronal holes, sunspots, and solar activity. *Progr. Phys.*, 2013, v. 2, L7–L9.
53. Robitaille P.-M. Commentary on the LMH model of the Sun II. Insight relative to coronal rain and splashdown events. *Progr. Phys.*, 2013, v. 2, L10–L11.
54. Robitaille P.M. Commentary on the LMH model of the Sun III. Insight into solar lithium abundances. *Progr. Phys.*, 2013, v. 2, L12–L13.
55. Robitaille P.-M. Commentary relative to the emission spectrum of the solar atmosphere: Further evidence for a distinct solar surface. *Progr. Phys.*, 2013, v. 3, L2–L4.
56. Robitaille P.-M. The LMH model of the Sun and the solar atmosphere I. Continuous emission and condensed matter within the chromosphere. *Progr. Phys.*, 2013, v. 3, L5–L7.
57. Robitaille P.-M. The LMH model of the Sun and the solar atmosphere II. Continuous emission and condensed matter within the corona. *Progr. Phys.*, 2013, v. 3, L8–L10.
58. Robitaille P.-M. The LMH model of the Sun and the solar atmosphere III. Importance of continuous emission spectra from flares, coronal mass ejections, prominences, and other coronal structures. *Progr. Phys.*, 2013, v. 3, L11–L14.
59. Robitaille P.-M. The LMH model of the Sun and the solar atmosphere IV. On the nature of the chromosphere. *Progr. Phys.*, 2013, v. 3, L15–L21.
60. Robitaille P.-M. The LMH model of the Sun and the solar atmosphere V. On the nature of the corona. *Progr. Phys.*, 2013, v. 3, L22–L25.
61. Robitaille P-M. The LMH model of the Sun and the solar atmosphere VI. Helium in the chromosphere. *Progr. Phys.*, 2013, v. 3, L26–L29.
62. Robitaille P.M. The LMH model of the Sun and the solar atmosphere VII. Further insights into the chromosphere and corona. *Progr. Phys.*, 2013, v. 3, L30–L36.
63. Agassi J. The Kirchhoff-Planck radiation law. *Science*, 1967, v. 156, no. 3771, 30–37.
64. Robitaille P.M. Max Karl Ernst Ludwig Planck (1858-1947). *Progr. Phys.*, 2007, v. 4, 117–120.
65. Heilbron J.L. The Dilemmas of an Upright Man — Max Planck and the Fortunes of German Science. Harvard University Press, Cambridge, MA, 1996.
66. Modest M.F. Radiative heat transfer. McGraw-Hill, New York, 1993, p. 92–108.
67. Carrington R.C. Observations on the Spots of the Sun, from November, 9, 1853, to March 24, 1861, Made at Redhill. Williams and Norgate, London, 1863.
68. Kitchatinov L.L. Solar differential rotation: Origin, models, and implications for dynamo. *ASI Conference Series — 1st Asia Pacific Solar Physics Meeting*, 2011, v. 2, 71–80 (also available online: arXiv:1108.1604v2 [astro-ph.SR] 9/4/2013).
69. Milne E.A. Thermodynamics of the stars. *Handbuch der Astrophysik*, 1930, v. 3, Part 1, 65–255 (also in Menzel D.H. Selected Papers on the Transfer of Radiation: 6 papers by Arthur Schuster, K. Schwarzschild, A.S. Eddington, S. Rosseland, E.A. Milne. Dover Publications, Inc., New York, 1966, p. 77–269).
70. Robitaille P.M. The little heat engine: Heat transfer in solids, liquids, and gases. *Progr. Phys.*, 2007, v. 4, 25–33.
71. Touloukian Y.S., DeWitt D.P. Thermophysical Properties of Matter (Vols. 1-8), IFI/Plenum, New York, 1972.
72. Thirumaleshwar M. Fundamentals of Heat and Mass Transfer, Dorling Kindersley, Dehli, 2009, p. 652.
73. Incropera F.P., DeWitt D.P., Bergman T.L. and Lavine A.S. Fundamentals of Heat and Mass Transfer, 6th Edition. John Wiley & Sons, Hoboken, NJ, 2007.
74. Langley S. P. Experimental determination of wave-lengths in the invisible prismatic spectrum. *Mem. Natl. Acad. Sci.*, 1883, v. 2, 147–162; Plate 12 and 21.
75. Langley S. P. On hitherto unrecognized wavelengths. *Phil. Mag.*, 1886, v. 22, 149–173; Plate IV–VI.
76. Kelly B.T. Physics of graphite. Applied Science Publishers, London, U.K., 1981, p. 34–61.
77. Delhaès P. World of carbon — vol. 1: Graphite and precursors. Gordon and Breach Science Publishers, Amsterdam, The Netherlands, 2001.
78. Pierson H.O. Handbook of carbon, graphite, diamond and fullerenes: Properties, processing and applications. Noyes Publications, Park Ridge, N.J., 1993.
79. Dresselhaus M.S. and Dresselhaus G. Intercalation compounds of graphite. *Adv. Phys.*, 2002, v. 1, no. 1, 1–186 (reprinted from *Adv. Phys.*, 1981, v. 30(2), 139–326).
80. Pietronero L. and Tosatti E. Physics of intercalation compounds. Springer-Verlag, Berlin, 1981.

81. Zabel H. and Solin S.A. Graphite intercalation compounds I: Structure and dynamics. Springer-Verlag, Berlin, 1990.
82. Dresselhaus M.S. and Kalish R. Ion implantation in diamond, graphite and related materials. Springer-Verlag, Berlin, 1992.
83. Enoki T., Suzuki M. and Endo M. Graphite intercalation compounds and applications. Oxford University Press, Oxford, U.K., 2003.
84. Martin W.H. and Brocklehurst J.E. The thermal expansion behavior of pyrolytic graphite-bromine residue compounds. *Carbon*, 1964, v. 1, no. 2, 133–141.
85. Hastings C.S. A theory of the constitution of the Sun, founded upon spectroscopic observations, original and other. *Am. J. Science*, 1881, v. 21, no. 121, 33–44.
86. Payne C.H. The relative abundances of the elements. Stellar Atmospheres. Harvard Observatory Monograph no. 1 (Harlow Shapley, Editor), Harvard University Press, Cambridge, MA, 1925 (reprinted in part in Lang K.R. and Gingerich O. A source book in astronomy and astrophysics, 1900–1975, Harvard University Press, Cambridge, MA, 1979, p. 245–248).
87. Russell H.N. On the composition of the Sun's atmosphere. *Astrophys. J.*, 1929, v. 70, 11–82.
88. Wigner E. and Huntington H.B. On the possibility of a metallic modification of hydrogen. *J. Chem. Phys.*, 1935, v. 3, 764–70.
89. Brovman E.G., Kagan Yu.M. and Kholas A. Properties of metallic hydrogen under pressure. *Sov. Phys. JETP*, 1972, v. 35, no. 4, 783–787.
90. Chakravarty S. and Ashcroft N.W. Ground state of metallic hydrogen. *Phys. Rev. B*, 1978, v. 18, 4588–4597.
91. Mon K.K., Chester G.V. and Ashcroft N.W. Simulation studies of a model of high-density metallic hydrogen. *Phys. Rev. B*, 1980, v. 21, 2641–2646.
92. McMahon J.M., Morales M.A., Pierleoni C. and Ceperley D.M. The properties of hydrogen and helium under extreme conditions. *Rev. Mod. Phys.*, 2012, v. 84, no. 4, 1607–1653.
93. Hubbard W.B., Guillot T., Lunime J.I., Burrows A., Saumon D., Marley M.S., Freedman R.S. Liquid metallic hydrogen and the structure of brown dwarfts and giant planets. *Phys. Plasmas*, 1997, v. 4, no. 5, 2011–2015 (also arXiv:astro-ph/9703007v1).
94. Nellis W.J., Ross M. and Holmes N.C. Temperature measurements of shock-compressed hydrogen: Implications for the interior of Jupiter. *Science*, 1995, v. 269, no. 5228, 1249–1252.
95. Nellis W.J., Weir S.T. and Mitchell A.C. Metallization and electrical conductivity of hydrogen in Jupiter. *Science*, 1996, v. 273, no. 5277, 936–938.
96. Lai D. and Salpeter E.E. Hydrogen phases on the surfaces of a strongly magnetized neutron star. *Astrophys. J.*, 1997, v. 491, 270–285.
97. Ichimaru S. Statistical Plasma Physics — Volume II: Condensed Plasmas, Addison-Westly, Redwood, CA, 1991 (reprinted by Westview Press, Boulder, CO, 2004).
98. Ichimaru S. and Kitamura H. Pycnonuclear reactions in dense astrophysical and fusion plasmas. *Phys. Plasmas*, 1999, v. 6, no. 7, 2649–2671.
99. Ichimaru S. Radiative proton-capture of high-Z nuclei in the sun and in liquid-metallic hydrogen. *Phys. Let. A*, 2000, v. 266, 167–172.
100. Scharmer G. B., Gudiksen B.V., Kiselman D., Löfdahl M.G., Rouppe van der Voort L.H.M. Dark cores in sunspot penumbral filaments. *Nature*, 2002, v. 420, 151–153.
101. Very F. The absorptive power of the solar atmosphere. *Astrophys. J.*, 1902, v. 16, 73–91.
102. Schwarzschild K. Über das Gleichgewicht der Sonnenamospäre. *Nachrichten von der Königlichen Gesellschaft der Wissenschaften zu Göttigen, Göttinger Nachrichten*, 1906, v. 195, 41–53 (English translation: Concerning the equilibrium of the solar atmosphere, found in: Menzel D.H. Selected Papers on the Transfer of Radiation: 6 papers by Arthur Schuster, K. Schwarzschild, A.S. Eddington, S. Rosseland, E.A. Milne. Dover Publications, Inc., New York, 1966, p. 25–34 and Meadows A.J. Early Solar Physics. Pergamon Press, Oxford, 1970, 277–290).
103. Milne E.A. Radiative equilibrium in the outer layers of a star: The temperature distribution and the law of darkening. *Mon. Not. Roy. Astron. Soc.*, 1921, v. 81, 361–375.
104. National Solar Observatory. Advanced Technology Solar Telescope — ATST. http://atst.nso.edu/sites/atst.nso.edu/files/press/ATST_book.pdf (accessed online on 9/10/2013).
105. Galilei G. and Scheiner C. On sunspots (translated by Eileen Reeves and Albert Van Helden), The University of Chicago Press, Chicago, IL, 2010.
106. Wilson A. Observations on the solar spots. *Phil. Trans. Roy. Soc.*, 1774, v. 64, 1–30.
107. Hale G.E. On the probable existence of a magnetic field in Sun-spots. *Astrophys. J.*, 1908, v. 28, 315–343.
108. Leonard T. and Choudhary D.P. Intensity and magnetic field distribution of sunspots. *Solar Physics*, 2008, v. 252, 33–41.
109. Martinez Pillet V. and Vázquez M. The continuum intensity-magnetic field relation in sunspot umbrae. *Astron. Astrophys.*, 1993, v. 270, 494–508.
110. Thomas J.H. and Weiss N.O. Sunspots and starspots. Cambridge University Press, Cambridge, U.K., 2008.
111. Sobotka M. Fine structure in sunspots. In, *Motions in the solar atmosphere* (A. Hanslmeier and M. Messerotti, eds.), Astrophysics and Space Science Library, v./,239, Kluwer Academic Publishers, Dordrecht, 1999, p. 71–97.
112. Langley S. Sur la température des diverses régions du soleil. Les noyaux noirs des taches. *Comptes Rendus*, 1875, v. 80, 746–749.
113. Frost E.B. Observations on the thermal absorption in the solar atmosphere made at Potsdam. *Astronnomische Nachrichten*, 1892, v. 130(3105–3106), 129–146.
114. Tandberg-Hanssen E. Solar Activity. Blaisdell Publishing Co., Waltham, M.A., 1967.
115. Bhatnagar A. Instrumentation and observational techniques in solar astronomy. In, *Lectures on Solar Physics* (H.M. Antia, A. Bhatnagar and R. Ulmschneider, Eds.), Springer, Berlin, 2003, p. 27–79.
116. Moradi H. and Cally P.S. Time-distance modeling in a simulated sunspot. *Solar Physics*, 2008, v. 251, 309–327.
117. Ilonidis S. and Zhao J. Determining absorption, emissivity reduction, and local suppression coeffcients inside sunspots. *Solar Physics*, 2011, v. 268, 377–388.
118. Bray R.J., Loughhead R.E., Durrant C.J. The solar granulation. Cambridge University Press, Cambridge, 1984.
119. Muller R. The solar granulation. In, *Motions in the solar atmosphere* (A. Hanslmeier and M. Messerotti, eds.), Astrophysics and Space Science Library, Kluwer Academic Publishers Dordrecht, 1999, v. 239, 35–70
120. Hirzberger J. Granulation and waves. *Astron. Nachr.*, 2003, v. 324, no. 4, 344–348.
121. Roudier Th. and Muller R. Structure of the solar granulation. *Solar Physics*, 1986, v. 107, 11–26.
122. Hirzberger J. On the brightness and velocity structure of solar granulation. *Astron. Astrophys.*, 2002, v. 392, 1105–1118.
123. Weaire D. and Rivier N. Soap, cells and statistics random patterns in two dimensions. *Contemp. Phys.*, 1984, v. 25, 59–99.
124. Rivier N. Structure of random cellular networks.*Science on Form: Proc. First Int. Symp. Sci. Form*, KTK Scientific Publishers, Tokyo, 1986, 451–458.

125. Chiu S.N. Aboav-Weaire's and Lewis' Law A review. *Mater. Charact.*, 1995, v. 34, 149–165.

126. Noever D.A. Solar granulation and statistical crystallography: a modeling approach using size-shape relations. *Astron. Astrophys.*, 1994, v. 282, 252–261.

127. Chacornac, On the physical constitution of the Sun. (25/12/1864). *The Reader*, 1865, v. 5, 16–17.

128. Rast M.P. On the nature of "exploding" granules and granule fragmentation. *Astrophys. J.*, 1995, v. 443, 863–868.

129. Hale G.E. The study of stellar evolution: An account of some recent methods of astrophysical research, The decennial publications of the University of Chicago — Second Series, Vol. X. University of Chicago Press, Chicago, 1908.

130. de la Rue W., Stewart B. and Loewy B. Researches on solar physics — Series II: On the behaviour of sun-spots with regard to increase and diminution (abstract). *Proc. Roy. Soc. London*, 1865, v. 14, 59–63.

131. Wang H., Deng N. and Liu C. Rapid transition of uncombed penumbrae to faculae during large flares. *Astrophys. J.*, 2012, v. 748, no. 2, 76(9 pp).

132. Chapman G.A. Facular line profiles and facular models. *Astrophys. J. Supp. Ser.*, 1977, v. 33, 35–54.

133. Tarbell T.D. and Title A.M. Measurements of magnetic fluxes and field strengths in the photospheric network. *Solar Physics*, 1977, v. 52, 13–25.

134. Ortiz A., Solanki S.K., Domingo V., Fligge M. and Sanahuja B. On the intensity contrast of solar photospheric faculae and network elements. *Astron. &Astrophys.*, 2002, v. 388, 1036–1047.

135. Libbrecht K.G. and Kuhn J.R. On the facular contrast near the solar limb. *Astrophys. J.*, 1985, v. 299, 1047–1050.

136. Spruit H.C. Pressure equilibrium and energy balance of small photospheric fluxtubes. *Solar Physics*, 1976, v. 50, 269–295.

137. Walton S.R. Flux tube models of solar plages. *Astrophys. J.*, 1987, v. 312, 909–929.

138. Bonet J.A., Cabello I. and Almeida S. Center-to-limb variation of the area covered by magnetic bright points in the quiet Sun. *Astron. & Astrophys.*, 2012, v. 539, A6.

139. Sanchez Almeida J., Bonet J.A., Viticchie B. and Del Moro D. Magnetic bright points in the quiet Sun. *Astrophys. J. Letters*, 2010, v. 715, no. 1, L26–L29.

140. Evershed J. Wave-length determinations and general results obtained from a detailed examination of spectra photographed at the solar eclipse of January 22, 1898. *Phil. Trans. Roy. Soc. London*, 1901, v. 197, 381–413.

141. Evershed J. Preliminary report of the expedition to the south limit of totality to obtain photographs of the flash spectrum in high solar latitudes. *Proc. Roy. Soc. London*, 1900, v. 67, 370–385.

142. Menzel D.H. and Cillié G.G. Hydrogen emission in the chromosphere. *Astrophys. J.*, 1937, v. 5, 88–106.

143. Athay R.G., Billings D.E., Evans J.W. and Roberts W.O. Emission in hydrogen Balmer lines and continuum in flash spectrum of 1952 total solar eclipse at Karthoum, Sudan. *Astrophys. J.*, 1954, v. 120, 94–111.

144. Athay R.G., Menzel D.H., Pecker J.C. and Thomas R.N. The thermo dynamic state of the outer solar atmosphere V. A model of the chromosphere from the continuous emission. *Astrophys. J. Suppl. Ser.*, 1955, v. 1, 505–519.

145. Hiei E. Continuous spectrum in the chromosphere. *Publ. Astron. Soc. Japan*, 1963, v. 15, 277–300.

146. Weart S.R. and Faller J.E. Photoelectric eclipse observation of the continuum at the extreme solar limb. *Astrophys. J.*, 1969, v. 157, 887–901.

147. Gingerich O. and de Jager C. The Bilderberg model of the photosphere and low chromosphere. *Solar Phys.*, 1968, v. 3, 5–25.

148. Ulmschneider P. The physics of the chromospheres and coronae. In, *Lectures on Solar Physics* (H.M. Antia, A. Bhatnagar and R. Ulmschneider, Eds.), Springer, Berlin, 2003, p. 232–280.

149. NASA. Air properties definitions. (Accessed online on 2/13/2013) www.grc.nasa.gov/WWW/k-12/airplane/airprop.html www.grc.nasa.gov/WWW/k-12/airplane/airprop.html (accessed online on 9/10/2013).

150. Bray R.J. and Loughhead R.E. The Solar Chromosphere, Chapman and Hall, London, U.K., 1974.

151. Dick S. Sky and Ocean Joined: The U.S. Naval Observatory 1830–2000. Cambridge University Press, Cambridge, 2003, p. 196–205.

152. Evershed J. Wave-length determinations and general results obtained from a detailed examination of spectra photographed at the solar eclipse of January 22, 1898. *Phil. Trans. Roy. Soc. London*, 1901, v. 197, 381–413.

153. Grotian W. Über die intensitätsverteilung des kontinuierlichen spektrums der inneren korona. *Zeitschrift fur Astrophysik*, 1931, v. 3, 199–226.

154. Lyot B. La couronne solaire étudiée en dehors des éclipses. *Comptes Rendus*, 1930, v. 191, 834–837.

155. Lyot B. A study of the solar corona and prominences without eclipses — George Darwin Lecture. *Mon. Not. Roy. Astron. Soc.*, 1939, 580-=594 (22 pages with plates).

156. Allen C.W. The spectrum of the corona at the eclipse of 1940 October 1. *Mon. Not. Roy. Astron. Soc.*, 1946, v. 106, 137–150.

157. Dwivedi B.N. The solar corona. In, *Lectures on Solar Physics* (H.M. Antia, A. Bhatnagar and R. Ulmschneider, Eds.), Springer, Berlin, 2003, p. 281–298.

158. van de Hulst H.C. The electron density of the corona. *Bull. Astronom. Inst. Netherlands*, 1950, v. 11, no. 410., 135–149.

159. Carrington R.C. Description of a singular appearance seen in the Sun on September 1, 1859. Mon. Not. Roy. Astron. Soc., 1859, v. 20, 13–15.

160. Machado M.E., Emslie A.G., Avrett E.H. Radiative backwarming in white-light flares. *Solar Phys.*, 1989, v. 124, 303–317.

161. Hiei E. A continuous spectrum of a white-light flare. *Solar Phys.*, 1982, v. 80, 113–127.

162. Wang H. Study of white-light flares observed by Hinode. *Res. Astron. Astrophys.*, 2009, v. 9, no. 2, 127–132.

163. Tandberg-Hanssen E. Solar prominences — An intriguing phenomenon. *Solar Phys.*, 2011, v. 269, 237–251.

164. Tandberg-Hanssen E. A spectroscopic study of quiescent prominences. *Astrophysica Novegica*, 1964, v. 9, no. 3, 13–32.

165. Yakovkin N.A. and Zeldina M.Yu. The prominence radiation theory. *Solar Phys.*, 1975, v. 45, 319–338.

166. Gopalswamy N. and Hanaoka Y. Coronal dimming associated with a giant prominence eruption. *Astrophys. J. Lett.*, 1998, v. 498, L179–L182.

167. Gopalswamy N., Shimojo M., Lu W., Yashiro S., Shibasaki K. and Howard R.A. Prominence eruptions and coronal mass ejections: A statistical study using microwave observations. *Astrophys. J.*, 2003, v. 586, 562–578.

168. Sheeley N.R., Walters J.H., Wang Y.M. and Howard R.A. Continuous tracking of coronal outflows: Two kinds of coronal mass ejections. *J. Geophys. Res.*, 1999, v. 104, no. A11, 24739–24767.

169. St. Cyr O.C., Howard R.A., Sheeley N.R., Plunkett S.P., Michels D.J., Paswaters S.E., Koomen M.J., Simnett G.M., Thompson B.J., Gurman J.B., Schwenn R., Webb D.F., Hildner E. and Lamy P.L. Properties of coronal mass ejections: SOHO LASCO observations from January 1996 to June 1998. *J. Geophys. Res.*, 2000, v. 105, no. A8, 18169–18185.

170. Hudson H.S. and Cliver E.W. Observing coronal mass ejections without coronagraphs. *J. Geophys. Res.*, 2001, v. 106, no. A11, 25199–25213.

171. Yashiro S., Gopalswamy N. Michaelek G, St. Cyr O.C., Plunkett S.P., Rich N.B. and Howard R.A. A catalog of white-light coronal mass ejections observed by the SOHO spacecraft. *J. Geophys. Res.*, 2004, v. 109, A07105(11 pages).

172. Wang Y.M. Nonradial coronal streamers. *Astrophys. J.*, 1996, v. 456, L119–L121.

173. Wang Y.M., Sheeley N.R., Walters J.H., Brueckner G.E., Howard R.A., Michels D.J., Lamy P.L., Schwenn R. and Simnett G.M. Origin of streamer material in the outer corona. *Astrophys. J.*, 1998, v. 498, L165–L168.

174. Vourlidas A. A review of white-light streamers at the end of cycle 23. *Proc. IAU: Solar Activity and its Magnetic Origin*, 2006, v. 233, 197–204.

175. Del Zanna G., Bromage B.J.I. and Mason H.E. Spectroscopic characteristics of polar plumes. *Astron. Astrophys.*, 2003, v. 398, 743–763.

176. Mouschovias T.Ch. and Poland A.I. Expansion and broadening of coronal loop transients: A theoretical explanation. *Astrophys. J.*, 1978, v. 220, 675–682.

177. Kjeldseth-Moe O. and Brekke P. Time variability of active region loops observed with the coronal diagnostic spectrometer (CDS) on SOHO. *Solar Phys.*, 1998, v. 182, 73–95.

178. Landi E., Miralles M.P., Curdt W. and Hara H. Physical properties of cooling plasma in quiescent active region loops. *Astrophys. J.*, 2009, v. 695, 221-237.

179. Ambstha A. The active and explosive Sun. In, *Lectures on Solar Physics* (H.M. Antia, A. Bhatnagar and R. Ulmschneider, Eds.), Springer, Berlin, 2003, p. 127–172.

180. Kahler S.W. Solar flares and coronal mass ejections. *Ann. Rev. Astron. Astrophys.*, 1992, v. 30, 113–141.

181. Priest E.R. Solar flare theory and the status of flare understanding. In *High Energy Solar Physics: Anticipating HESSI. ASP Conf. Ser.*, 2000, v. 206, 13–26.

182. Priest E.R. and Forbes T.G. The magnetic nature of solar flares. *Astron. Astrophys. Rev.*, 2002, v. 10, 313–377.

183. Hudson H.S. Global properties of solar flares. *Space Sci. Rev.*, 2011, v. 158, 5–41.

184. Holman G.D. The mysterious origin of solar flares. *Sci. Am.*, 2006, v. 294, no. 4, 38–45.

185. Forbes T.G. Magnetic reconnection in solar flares. *Geophys. Astrophys. Fluid Dynam.*, 1991, v. 62, 15–36.

186. Ramaty R., Mandzhavidze N., Kozlovsky B. and Murphy R.J. Solar atmospheric abundances and energy content in flare-accelerated ions from gamma-ray spectroscopy. *Astrophys. J.*, 1995, v. 455, L193–L196.

187. Watanabe K., Krucker S., Hudson H., Shimizu T., Masuda S., and Ichimoto K. G-band and hard X-ray emissions of the 2006 December 14 flares observed by HINODE/SOT and RHESSI. *Astrophys. J.*, 2010, v. 715, 651–655.

188. Anderson J.L. Hinode discovers origin of white-light flares. NASA–Marshall Space Flight Center. Accessed online on 4/17/2013: http://www.nasa.gov/centers/marshall/news/news/releases/2010/10-052.html

189. Zöllner F. On the temperature and physical constitution of the Sun. *Phil. Mag. 4th Series*, 1870, v. 40, 313–327 (essentially reprinted in: Zöllner F. On the Sun's temperature and physical constitution. *Nature*, 1870, v. 2(52), 522–526).

190. Kirchhoff G. Researches on the Solar Spectrum, and the Spectra of the Chemical Elements (Translated by H.E. Roscoe), Macmillan and Co., Cambridge, U.K., 1862.

191. Kirchhoff G. and Bunsen R. Chemical analysis by spectral observations. (Reprinted from *Poggendorf's Annalen der Physik*, 1860, v. 110, In, *The laws of Radiation and Absorption — Memoirs by Prévost, Stewart, Kirchhoff, and Kirchhoff and Bunsen.* (D.B. Brace, Ed.), American Book Company, N.Y., 1901, p. 99–126).

192. Phillips K.J.H., Feldman U. and Landi E. Ultraviolet and X-Ray Spectroscopy of the Solar Atmosphere: Cambridge Astrophysics Series — Vol. 44, Cambridge University Press, Cambridge, U.K., 2008.

193. Zirin H. The mystery of the chromosphere. *Solar Phys.*, 1996, v. 169, 313–326.

194. Rieger E. Solar flares: High-energy radiation and particles. *Solar Phys.*, 1989, v. 121, 323–345.

195. Ramaty R. and Simnett G.M. Accelerated particles in solar flares. In, *The Sun in Time* (C.P. Sonett, M.S. Giampapa and M.S. Matthews, Eds.), The University of Arizona Press, Tuscon, AZ, 1991, p. 232–259.

196. Grevesse N. Solar abundances of lithium, beryllium, and boron. *Solar Phys.*, 1968, v. 5, 159–180.

197. Wiens R.C., Bochsler P., Burnett D.S. and Wimmer-Schweingruber R.F. Solar and solar-wind isotopic compositions. *Earth Plan. Sci. Let.*, 2004, v. 222, 697–712.

198. Uitenbroek H. The effect of photospheric granulation on the determination of the lithium abundance in solar-type stars. *Astrophys. J.*, 1998, v. 498, 427–440.

199. Schatzman E. Turbulent transport and lithium destruction in main sequence stars. *Astron. Astrophys.*, 1977, v. 56, 211–218.

200. Pinsonneault M.H. Rotational mixing and lithium depletion. In, *Light Elements in the Universe: Proc. IAU Symposium*, 2009, v. 268, 375-380.

201. Pinsonneault M.H. A fossil record for exoplanets. *Nature*, 2009, v. 462, no. 7270, 168–169.

202. Israelian G., Delgado Mena E., Santos N.C., Sousa S.G., Mayor M., Udry S., Dominguez Cerdena C., Rebolo R. and Randich S. Enhanced lithium depletion in Sun-like stars with orbiting planets. *Nature*, 2009, v. 462, no. 7270, 189–191.

203. Baumann P., Ramirez I., Melendez J., Aslund M. and Lind K. Lithium depletion in solar-like stars: no planet connection. *Astron. Astrophys.*, 2010, v. 519, A87(11 pages).

204. Zurek E., Hoffmann R., Ashcroft N.W., Oganov A.R., Lyakhov A.O. A little bit of lithium does a lot for hydrogen. *Proc. Nat. Acad. Sci. USA* USA, 2009, v. 106, no. 42, 17640—17643.

205. Menzel D.H. A Study of the Solar Chromosphere. *Publications of the Lick Observatory*, University of California Press, Berkeley, CA, v. 17, 1931.

206. Athay R.G. The Solar Chromosphere and Corona: Quiet Sun. D. Reidel Publishing Co., Boston, M.A., 1976.

207. Thomas R.N. and Athay R.G. Physics of the Solar Chromosphere. Interscience Publishers, New York, N.Y., 1961.

208. Heinzel P. Understanding the solar chromosphere. In, *"Exploring the Solar System and the Universe"*, (V. Mioc, C. Dumitrache, N.A. Popescu, Eds.), American Institute of Physics, 2008, 238–244.

209. Woltjer L. A photometric investigation of the spicules and the structure of the chromosphere. *Bull. Astron. Inst. Netherlands*, 1954, v. 12, no. 454, 165–176.

210. Rush J.H. and Roberts W.O. Recent studies of chromospheric spicules. *Australian J. Phys.*, 1954, v. 7, 230–243.

211. Beckers J.M. Solar spicules. *Ann. Rev. Astron. Astrophys.*, 1972, v. 10, 73–100.

212. Lorrain P. and Koutchmy S. Two dynamical models for solar spicules. *Solar Phys.*, 1996, v. 165, 115–137.

213. Sterling A. Solar spicules: A review of recent models and targets for future observations. *Solar Phys.*, 2000, v. 196, 79–111.

214. Zaqarashvili T.V. and Erdélyi R. Oscillations and waves in solar spicules. *Space Sci. Rev.*, 2009, v. 149, 355–388.

215. Pasachoff J.M., Jacobson W.A. and Sterling A.C. Limb spicules from ground and from space. *Solar Phys.*, 2009, v. 260, 59–82.

216. Konig L., Rabin I, Schultze W. and Ertl G. Chemiluminescence in the agglomeration of metal clusters. *Science*, 1996, v. 274, no. 5291, 1353–1355.

217. Jordan C., Bueckner G.E., Bartoe J.D.F., Sandlin G.D. and Vanhoosier M.E. Emission lines of H2 in the extreme-ultraviolet solar spectrum. *Astrophys. J.*, 1978, v. 226, 687–697.

218. Bartoe J.D.F., Brueckner G.E., Nicolas K.R., Sandlin G.D., Vanhoosier M.E. and Jordan C. H2 emission in the solar atmosphere. *Mon. Not. Roy. Astron. Soc.*, 1979, v. 187, 463–471.

219. Sandlin G.D., Bartoe J.D.F., Brueckner G.E., Tousey R. and Vanhoosier M.E. The high-resolution solar spectrum, 1175-1770Å. *Astrophys. J. Suppl. Ser.*, 1986, v. 61, 801–898.

220. Innes D.E. SUMER-Hinode observations of microflares: excitation of molecular hydrogen. *Astron. Astrophys.*, 2008, v. 481, no. 1, L41–L44.

221. Przybilla N. and Butler K. The solar hydrogen spectrum in non-local thermodynamic equilibrium. *Astrophys. J.*, 2004, v. 610, L61–L64.

222. Kurucz R. ATLAS9 Stellar Atmosphere Programs and 2 km/s grid — CD-ROM No. 13, *Smithsonian Astrophysical Observatory*, 1993.

223. Stehlé C. and Hutcheon R. Extensive tabulations of Stark broadened hydrogen line profiles. *Astro. & Astrophys. Suppl. Ser.*, 1999, v. 140, 93–97.

224. Konjević N., Lesage A., Fuhr J.R. and Wiese W.L. Experimental Stark widths and shifts for spectral lines of neutral and ionized atoms (A critical review of selected data for the period 1989 through 2000). *J. Phys. Chem. Ref. Data*, 2002, v. 31, no. 3, 819–927.

225. Stehlé C. Stark broadening of hydrogen Lyman and Balmer lines in the conditions of stellar envelopes. *Astron. Astrophys. Suppl. Series*, 1994, v. 104, 509–527.

226. Stehlé C. and Jacquemot S. Line shapes in hydrogen opacities. *Astron. Astrophys.*, 1993, v. 271, 348–359.

227. Acon B.W., Stehlé C., Zhang H. and Montaser A. Stark-broadened hydrogen line profiles predicted by the model microfield method for calculating electron densities. *Spectrochimica Acta B: Atomic Spectros.*, 2001, v. 56, 527–539.

228. Mijatović Z., Konjević N., Ivković M. and Kobilarov R. Influence of ion dynamics on the width and shift of isolated He I lines in plasmas — II. *Phys. Rev. E*, 1995, v. 51, no. 5, 4891–4896.

229. Cairns I.H., Lobzin V.V., Warmuth A., Li B., Robinson P.A., and Mann G. Direct radio probing and interpretation of the Sun's plasma density profile. *Astrophys. J. Letters*, 2009, v. 706, L265–L269.

230. Wöhl H. On molecules in sunspots. *Solar Phys.*, 1971, v. 16, 362–372.

231. Sinha K. Molecules in the Sun. *Proc. Astron. Soc. Australia*, 1991, v. 9, 32–36.

232. Olmsted C.M. Sun-spot bands which appear in the spectrum of a calcium arc burning in the presence of hydrogen. *Astrophys. J.*, 1908, v. 27, 66–69.

233. Ionization Potentials. — Accessed on 10/5/2013. http://www.physics.ohio-state.edu/~lvw/handyinfo/ips.html.

234. Athay R.G. Line broadening in chromospheric spicules. *Astrophys. J.*, 1961, v. 134, 756–765.

235. Zirker J.B. The solar H and K lines of ionized calcium. *Solar Phys.*, 1968, v. 3, 164–180.

236. Gulyaev R.A. and Livshits M.A. Width of Ca^+H line in spicules. *Soviet Astron.*, 1966, v. 9, no. 4, 661–663.

237. Janssen J. Indications de quelques-uns des résultats obténus à Guntoor, pendant l' éclipse du mois d'août dernier. *Comptes Rendus*, 1868, v. 67, 838–39.

238. Lockyer J.N. Notice of an observation of the spectrum of a solar prominence. *Proc. Roy. Soc. London*, 1868, v. 17, 91–92.

239. Kragh H. The solar element: A reconsideration of helium's early history. *Ann. Science*, 2009, v. 66, no. 2, 157–182.

240. Taylor R.J. The helium problem. *Quart. J. Roy. Astron. Soc.*, 1967, v. 8, 313–333.

241. Grevesse N. and Sauval A.J. Standard solar composition. *Space Science Reviews*, 1998, v. 85, 161–174.

242. Asplund M., Grevesse N., Sauval A.J. and Scott P. The chemical composition of the Sun. *Ann. Rev. Astron. Astrophys.*, 2009, v. 47, 481–522.

243. Auchère F., Boulade S., Koutchmy S, Smartt R.N., Delaboudinière J.P., Georgakilas A., Gurman J.B. and Artzner G.E. The prolate solar chromosphere. *Astron. & Astrophys.*, 1998, v. 336, L57–L60.

244. Marchenko G.P., Akimov L.A., Belkina I.L. and Velikodsky Y.I. Solar chromosphere in the D_3 helium line from spectra of the eclipse of March 29, 2006. *Kinematics Phys. Celestial Bodies*, 2011, v. 27, no. 4, 183–190.

245. Zirin H. The Solar Atmosphere. Cambridge University Press, Cambridge, U.K., 1998.

246. Centeno R., Bueno J.T., Uitenbroek H. and Collados M. The influence of coronal EUV irradiance on the emission in the HeI 10830Å and D_3 multiplets. *Astrophys. J.*, 2008, v. 677, 742–750.

247. Grandinetti F. Helium chemistry: A survey of the role of the ionic species. *Inter. J. Mass Spectrom.*, 2004, v. 237, 243–267.

248. Hogness T.R. and Lunn E.G. The ionization of hydrogen by electron impact as interpreted by positive ray analysis. *Phys. Rev.*, 1925, v. 26, 44–55.

249. Tolliver D.E., Kyrala G.A and Wing W. H. Observation of the infrared spectrum of helium-hydride molecular ion $^4HeH^+$. *Phys. Rev. Let.*, 1979, v. 43, no. 23, 1719–1722.

250. Crofton M.W., Altman R.S., Haese N.N. and Oka T. Infrared spectra of $^4HeH^+$, $^4HeD^+$, $^3HeH^+$, and $^3HeD^+$. *J. Chem. Phys.*, 1989, v. 91, 5882-5886.

251. Roberge W. and Dalgarno A. The formation and destruction of HeH^+ in astrophysical plasmas. *Astrophys. J.*, 1982, v. 255, 489–496.

252. Engel E.A., Doss N., Harris G.J. and Tennyson J. Calculated spectra for HeH^+ and its effect on the opacity of cool metal-poor stars. *Mon. Not. Roy. Astron. Soc.*, 2005, v. 357, 471–477.

253. Galli D. and Palla F. The chemistry of the early universe. *Astron. Astrophys.*, 1998, v. 335, 403–420.

254. Ketterle W., Figger H. and Walther H. Emission spectra of bound helium hydride. *Phys. Rev. Lett.*, 1985, v. 55, no. 27, 2941–2944.

255. www.nobelprize.org/nobel_prizes/physics/laureates/2001/ketterle.html

256. Ketterle W., Figger H. and Walther H. Emission spectra of bound helium hydride. *Phys. Rev. Lett*, 1985, v. 55, 2941–2944.

257. Ketterle W., Dodhy A. and Walther H. Bound-free emission of the helium hydride molecule. *Chem. Phys. Let.*, 1986, v. 129, no. 1, 76–78.

258. Wollaston W.H. A method of examining refractive and dispersive powers, by prismatic reflection. *Phil. Trans. Roy. Soc. London*, 1802, v. 92, 365–380.

259. Fraunhofer J. Bestimmung des Brechungs-und des Farbenzerstreuungs-Vermögens verschiedener Glasarten, in Bezug auf die Vervollkommnung achromatischer Fernröhre. *Denkschriften der Königlichen Akademie der Wissenschaften zu München*, 1814/1815, v. 5, 193–226.

260. Unsöld A. Über die Struktur der Fraunhofersehen Linien und die quantitative Spektralanalyse der Sonnenatmosphäre. Zeitschrift für Physik, 1928, v. 46, no. 11–12, 765–781.

261. Kuli-Zade D.M. Profiles of the stronger Fraunhofer lines in the solar spectrum. I. Hydrogen lines. *Sov. Astron.*, 1965, v. 8, no. 5, 736–743.

262. Bondal K.R. and Gaur V.P. On some visible and infrared atomic hydrogen lines in three photospheric models. *Bull. Astron. Soc. India*, 1986, v. 14, 73–78.

263. de Jager C., Migeotte M. and Neven L. The profile of the Brackett alpha line in the solar spectrum. *Ann. Astrophys.*, 1956, v. 19, 9–18.

264. Cowley C.R. and Castelli F. Some aspects of the calculation of Balmer lines in the Sun and stars. *Astron. & Astrophys.*, 2002, v. 387, 595–604.

265. Hufbauer K. Exploring the Sun: Solar Science since Galileo. The Johns Hopkins University Press, Baltimore, 1991, p. 112–114.

266. Claridge G.C. Coronium. *J. Roy. Astron. Soc. Canada*, 1937, v. 31, no. 8, 337–346.

267. Unsigned. Origin of the coronium lines. *Nature*, 1942, v. 150, no. 3817, 756–759.

268. Milne A.E. Presidential Address — Award of the Gold Medal to Professor Bengt Edlén. *Mon. Not. Roy. Astron. Soc.*, 1945, v. 105, 138–145.

269. Wood B.E., Karovska M., Cook J.W., Brueckner G.E., Howard R.A., Korendyke C.M. and Socker D.G. Search for brightness variations in FeXIV coronagraph observations of the quiescent solar corona. *Astrophys. J.*, 1998, v. 505, 432–442.

270. Habbal S.R., Druckmüller M., Morgan H., Daw A., Johnson J., Ding A., Arndt M., Esser R., Rušin V. and Scholl I. Mapping the distribution of electron temperature and Fe charge states in the corona with total solar eclipse observations. *Astrophys. J.*, 2010, v. 708, 1650–1662.

271. Habbal S.R., Druckmüller M., Morgan H., Scholl I., Rušin V., Daw A., Johnson J. and Arndt M. Total solar eclipse observations of hot prominence shrouds. *Astrophys. J.*, 2010, v. 719, 1362–1369.

272. Habbal S.R., Morgan H. and Druckmüller M. A new view of coronal structures: Implications for the source and acceleration of the solar wind — First Asia-Pacific Solar Physics Meeting. *ASI Conf. Ser.*, 2011, v. 2, 259–269.

273. Priest E.R. Solar flare theory and the status of flare understanding. *High Energy Solar Physics: Anticipating HESSI (R. Ramaty and N. Mandzhavidze, Eds.), ASP Conf. Ser.*, 2000, v. 206, 13–26.

274. Rakov V.A. and Uman M.A. Lightning: Physics and Effects. Cambridge University Press, Cambridge, U.K., 2003.

275. Uman M.A. Lightning. Dover Publicatons. New York, N.Y., 1984.

276. Uman M.A. The Lightning Discharge (International Geophysics Series — Vol. 39), Academic Press, Inc., New York, N.Y., 1987.

277. Bhatnagar A. and Livingston W. Fundamentals of Solar Astronomy (World Scientific Series in Astronomy and Astrophysics — Vol. 6), World Scientific, New Jersey, 2005.

278. Jeans J.H. The Universe Around Us. 1st Edition, Cambridge University Press, 1933.

279. NASA. Factsheets. (Accessed online on 9/14/13)
http://nssdc.gsfc.nasa.gov/planetary/factsheet/earthfact.html
http://nssdc.gsfc.nasa.gov/planetary/factsheet/jupiterfact.html
http://nssdc.gsfc.nasa.gov/planetary/factsheet/saturnfact.html
http://nssdc.gsfc.nasa.gov/planetary/factsheet/neptunefact.html
http://nssdc.gsfc.nasa.gov/planetary/factsheet/uranusfact.html

280. Faye H. Sur la constitution physique du Soleil. Les Mondes, 1865, v. 7, 293–306 (translated into English by Patrice Robitaille: On the physical constitution of the Sun — Part I. *Progr. Phys.*, 2011, v. 3, 35–40.

281. Emilio M., Kuhn J.R., Bush R.I, and Scholl I.F. Measuring the solar radius from space during the 2003 and 2006 Mercury transits. *Astrophys. J.*, 2012, v. 750, 135(8 pages).

282. Chapman G.A., Dobias J.J., and Walton S.R. On the variability of the apparent solar radius. *Astrophys. J.*, 2008, v. 681, no. 2, 1698–1702.

283. Rozelot J.P. and Damiani C. Rights and wrongs of the temporal solar radius variability. *Eur. Physical J. H*, 2012, v. 37, no. 5, 709–743.

284. Chandrasekhar S. Ellipsoidal Figures of Equilibrium. Yale Univerity Press, New Haven, 1969.

285. Domiciano de Souza A., Kervella P., Jankov S., Abe L., Vakili F., di Folco E. and Paresce F. The spinning-top Be star Achernar from VLTI-VINCI. *Astron. & Astrophys.*, 2003, v. 407, L47–L50.

286. Carciofi A.C., Domiciano de Souza A., Magalhães A.M., Bjorkman J.E. and Vakili F. On the Determination of the Rotational Oblateness of Achernar. *Astrophys. J.*, v. 676, no. 1, L41–L44.

287. Godier S and Rozelot J.P. The solar oblateness and its relationship with the structure of the tacholine and the Sun's subsurface. *Astron. & Astrophys.*, 2000, v. 355, 365–374.

288. Gough D. How oblate is the Sun? *Science*, 2012, v. 337, 1611–1612.

289. Kuhn J.R., Bush R., Emilio M. and Scholi I.F. The precise solar shape and its variability. *Science*, 2012, v. 337, 1638–1640.

290. Rozelot J.P., Damiani C. and Lefebvre S. Variability of the solar shape (before space dedicated missions). *J. Atmospher. Solar-Terrestrial Phys.*, 2009, v. 71, 1683–1694.

291. Roberts A.M. Solar faculae stand exposed. *Sky & Telescope*, 2003, v. 106, no. 4, 26.

292. Scharmer G.B., Henriques V.M.J., Kiselman D. and de la Cruz Rodríguez. Detection of convective downflows in a sunspot umbra. *Science*, 2011, v. 333, no. 6040, 316–319.

293. AAAS. AAAS News Release — "SCIENCE: Pinpointing the Nature of Filaments in Sunspots." www.aaas.org, 7 Sept 2011

294. Thomas J.H. Solar physics: The Sun under a microscope. *Nature*, 2002, v. 420, 134–135.

295. Cranmer S.R. Coronal holes. In, *Encyclopedia of Astronomy and Astrophysics*, (Paul Murdin, Ed.), Institute of Physics Publishing, Bristol, 2001, v. 1, p. 496–501.

296. Cranmer S.R. Coronal holes. *Living Rev. Solar Phys.*, 2009, v. 6, 3–66.

297. Chitre S.M. Overview of solar physics. In, *Lectures on Solar Physics* (H.M. Antia, A. Bhatnagar and R. Ulmschneider, Eds.), Springer, Berlin, 2003, p. 1–26.

298. Lewis D.J., Simnett G.M., Brueckner G.E., Howard R.A., Lamy P.L. and Schwenn R. LASCO observations of the coronal rotation. *Solar Phys.*, 1999, v. 184, no. 2, 297–315.

299. Shevick E. Science Action Labs: Astronomy (Illustrated by Rojas M.G.), Teaching & Learning Company, Dayton, OH, 2002.

300. Moore P. and Watson J. Observing the Solar System. In *Astronomy with a Budget Telescope, Patrick Moore's Practical Astronomy Series*, 2012, p. 31-64.

301. Jenkins J.L. The Sun, Yesterday and Today in *The Sun and How to Observe it: Astronomers' Observing Guides*, Springer, New York, 2009, p. 1–15.

302. University Corporation for Atmospheric Research. A closer look. Approved by Peter Fox Last revised: Mon Apr 10 15:08:11 MDT 2000. http://www.ucar.edu/communications/lasers/sun/closer.html (accessed on 9/10/2013)

303. NASA, Release No. 67-124, Mariner E/Venus 67, 1967 p. 20. http://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/19670017987_1967017987.pdf (accessed on 9/10/2013)

304. Saarman E. Sun seeker, Stanford Report, May 10, 2006. http://news.stanford.edu/news/2006/may10/hoeksema-051006.html (accessed on 9/10/2013)

305. NASA, Astronomy Picture of the Day, January 10, 2000. http://apod.nasa.gov/apod/ap000110.html (accessed on 9/16/2013)

306. L.A. Times. New images show Sun is covered by wide mountains of hot gases. December 16, 1996. http://articles.latimes.com/1996-12-19/local/me-10771_1_hot-gases (accessed on 9/10/2013)

307. Lang K. SOHO reveals the secrets of the Sun, *Scientific Am.*, 1997, v. 276(3), p. 40(8 pages).

308. Ludwig G.H. Particles and fields in space research. *Geophys. Astron. Space Exploration NASA SP-13*, 1962, p.21–29.

309. Baker D.N. Critical issues in space plasma physics. *Phys. Plasmas* 1999, 6(5), 1700–1708.

310. Jordan S.D. The solar optical telescope (SOT), *Space Science Rev.*, 1981, v. 29(4), 333-340.

311. Marhavilas P.K. The space as a natural laboratory of electrotechnics. *J. Eng. Science Tech. Rev.*, 2008, v. 1, 9–18.

312. Verschuur G. The Radio Sun and Planets in *The Invisible Universe: The Story of Radio Astronomy*, Springer, New York, 2007, p. 23–32.

313. Filipov B. and Koutchmy S. On the origin of the prolate solar chromosphere. *Solar Phys.*, 2000, v.196, 311–320.

314. Bénard H. Les tourbillons céllulaires dans une nappe liquide. *Rev. Gen. Sci. Pures Appl.*, 1900, v. 11, 1261–1271; 1309–1328.

315. Bénard H. Les toubillons céllulaires dans une nappe liquid transportant de la chaleur par convection en régime permanent. *Ann. Chim.Phys.*, 1901, v. 23, 62–144.

316. Ozbelge H.O. On the surface structure and the hydrodynamics of the Bénard Cells. *Experiment. Fluids*, 1989, v. 8, 238–240.

317. Cerisier P., Rahal S., Rivier N. Topological correlations in Bénard-Marangoni convective structures. *Phys. Rev. E.*, 1996, v. 54, no. 5, 5086–5094.

318. Thiele U. and Eckert K. Stochastic geometry of polygonal networks an alternative approach to the hexagon-square-transition in Bénard convection. *Phys. Rev. E*, 1998, v. 58, 3458–3468

319. Charbonneau P. and Smolarkiewicz P.K. Modeling the Solar Dynamo. *Science*, 2013, v. 340, no. 6128, 42–43.

320. Tobias S.M. The Solar Dynamo. *Phil. Trans. Roy. Soc. A*, 2002, v. 360, no. 1801, 2741–2756.

321. Nornberg M.D., Spence E.J., Kendrick R.D., Jacobson C.M. and Forest C.B. Measurements of the magnetic field induced by a turbulent flow of liquid metal. *Phys. Plasmas*, 2006, v. 13, 055901(8 pages); also arXiv:physics/0510265v3 [physics.plasm-ph].

322. Nornberg M.D., Spence E.J., Kendrick R.D., Jacobson C.M. and Forest C.B. Intermittent Magnetic Field Excitation by a Turbulent Flow of Liquid Sodium. *Phys. Rev. Let.*, 2006, v. 97, no.4, 044503(4 pages).

323. Stieglitz R. and Müller U. Experimental demonstration of a homogeneous two-scale dynamo. *Phys. Fluids*, 2001, v. 13, no. 3, 561–564.

324. Gailitis A., Lielausis O., Dement'ev S., Platacis E., Cifersons A., Gerbeth G., Gundrum T., Stefani F., Christen M., Hänel H. and Will G. Detection of a Flow Induced Magnetic Field Eigenmode in the Riga Dynamo Facility. *Phys. Rev. Let.*, 2000, v. 84, no. 19, 4365–4368.

325. Density of molten elements and representative salts. In, *CRC Handbook of Chemistry and Physics*, 2013–2014, CRC Press, Baca Raton, FL, p. 4-128.

326. Schecter D.A., Boyd J.F. and Gilman P.A. "Shallow-Water" Magnetohydrodynamic Waves in the Solar Tachocline. *Astrophys. J.*, v. 551, L185–L188.

327. Mohorič A., Planinšič G., Kos M., Duh A. and Stepišnik J. Magnetic Resonance Imaging System Based on Earth's Magnetic Field. *Inst. Science Tech.*, 2004, v./,32, no. 6, 655–667.

328. Lowes F.J. and Wilkinson I. Geomagnetic dynamo: A laboratory model. *Nature*, 1963, v. 198, no. 4886, 1158–1160.

329. Asinovskii E.I. and Markovets V.V. The limiting electric conductivity of plasma. *Phys. Letters A*, 2003, v. 319, 510–513.

330. De Groof A., Berghmans D., van Driel-Gesztelyi L. and Poedts S. Intensity variations in EIT shutterless mode: Waves or Flows? *Astron. Astrophys.*, 2004, v. 415, no. 3, 1141–1151.

331. De Groof A., Bastiaensen C., Müller D.A.N., Berghmans D., and Poedts S. Detailed comparison of downflows seen both in EIT 30.4 nm and Big Bear Hα movies. *Astron. Astrophys.*, 2005, v. 443, no. 1, 319–328.

332. Müller D.A.N., De Groof A., Hansteen V.H. and Peter H. High-speed coronal rain. *Astron. Astrophys.*, 2005, v. 436, no. 3, 1067–1074.

333. Antolin R., Vissers G. and van der Voort L.R. On-Disk coronal rain. *Solar Phys.*, 2012, v. 280, no. 2, 457–474.

334. Antolin R. and van der Voort L.R. Observing the fine structure of loops through high-resolution spectroscopic observations of coronal rain with the CRISP instrument at the Swedish Solar Telescope. *Astrophys. J.*, 2012, v. 745, no. 2, 152(21 pages).

335. Aschwanden M. Physics of the Solar Corona: An Introduction with Problems and Solutions. Springer in association with Praxis Publishing, Chichester, U.K., 2005.

336. Peter H., Bingert S., Klimchuk J.A., de Forest C., Cirtain J.W., Golub L., Winebarger A.R., Kobayashi K. and Korreck K.E. Structure of solar coronal loops: From miniature to large-scale. *Astron. & Astrophys.*, 2013, v. 556, A104(12 pages).

337. Tsiropoula G., Tziotsious K., Kontogiannis I., Madjarska M.S., Doyle J.G. and Suematsu Y. Solar fine-scale structures I. Spicules and other small-scale, jet-like events at the chromospheric level: Observations and physical parameters. *Space Sci. Rev.*, 2012, v. 169, 181–244.

338. De Pontieu B., McIntosh S.W., Hansteen V.H., Carlsson M., Schrijver C.J., Tarbell T.D., Title A.M., Shine R.A., Suematsu Y., Tsuneta S., Katsukawa Y., Ichimoto K., Shimizu T. and Nagata S. A tale of two spicules: The impact of spicules on the magnetic chromosphere. *Publ. Astron. Soc. Japan*, 2007, v. 59, 655–660.

339. Zhang Y.Z., Shirata K, Wang J.X., Mao X.J., Matsumoto T., Liu Y. and Su T.T. Revision of spicule classification. *Astrophys. J.*, 2012, v. 750, 16(9 pages).

340. Moskowitz C. Gargantuan Sun explosion rocks astronomers. *SPACE.com* (created June 9, 2011 at 12:34 PM ET — accessed online on January 29, 2011).

341. NASA/SDO/Helioviewer.org [2011/06/07 04:00:00 to 11:00:00 UTC]. Observed well using 5 min frames SDO AIA 304. These events have been captured in video format and posted numerous times online: e.g. Newsflash *Skywatch Media* YouTube Channel; Phil Plait *Bad Astronomy Blog* on YouTube. (Accessed online on January 29, 2013).

342. Formisano V. and Moreno G. Helium and heavy ions in the solar winds. *Revista del Nuovo Cimento*, 1971, v. 1, no. 3, 365–422.

343. Robbins D.E., Hundhausen A.J. and Bame S.J. Helium in the solar-wind. *J. Geophys. Res.*, 1970, v. 75, no. 7, 1178–1187.

344. Ogilvie K.W. and Hirshberg J. The solar cycle variation of the solar wind helium abundance. *J. Geophys. Res.*, 1974, v. 79, no. 31, 4595–4602.

345. Bame S.J., Asbridge J.R., Feldman W.C. and Gosling J.T. Evidence for a structure-free state at high solar wind speeds. *J. Geophys. Res.*, 1977,v. 82, no. 10, 1487–1492.

346. Borrini G., Gosling J.T., Bame S.J., and Feldman W.C. Helium abundance enhancements in the solar wind. *J. Geophys. Res.*, 1982, v. 87, no. A9, 7370–7378.

347. Bochsler P. Structure of the solar wind and compositional variations. *Space Sci. Rev.*, 1998, v. 85, 291–302.

348. Aellig M.R., Lazarus A.J. and Steinberg J.T. The solar wind helium abundance: Variations with wind speed and solar cycle. *Geophys. Res. Let.*, 2001, v. 28, no. 14, 2767–2770.

349. Wurz P. Solar Wind Composition. In: The Dynamic Sun: Challenges for Theory and Observations, *ESA SP-600*, 2005, v. 5.2, 1–9.

350. Kasper J.C., Stevens M.L., Lazarus A.J. and Ogilvie K.W. Solar wind and helium abundance as a function of speed and heliographic latitude: Variation through a solar cycle. *Astrophys. J.*, 2007, v. 660, 901–910.

351. Wang Y.M. Relating the solar wind helium abundance to the coronal magnetic field. *Astrophys. J.*, 2008, v. 683, 499–509.

352. Parker E.N. Dynamics of the interplanetary gas and magnetic fields. *Astrophys. J.*, 1958, v. 128, 664–675.

353. Biermann L. Über die Ursache der chromosphärischen Turbulenz und des UV-Exzesses der Sonnenstrahlung. *Zeitshrift für Physik*, 1948, v. 25, 161–177.

354. Unsöld A. and Chapman S. Optical and radiofrequency ansorption by solar corpuscular burts. *The Observatory*, 1949, v. 69, 219–221.

355. Chapman S. The Viscosity and Thermal Conductivity of a Completely Ionized Gas. *Astrophys. J.*, 1954, v. 120, 151–155.

356. Gough D.O., Leibacher J.W., Scherrer P.H., Toomre J. Perspectives in Helioseismology. *Science*, 1996, v. 272, no. 5266, 1281–4.

357. Gough D.O. Seismology of the Sun and distant stars. D. Riedel Publishing Company. Dordrecht, 1986.

358. Kosovichev A.G. Solar Oscillations. In: Stellar Pulsation — Challenges for Theory and Observation. *AIP Int. Conf. Series.*, 2009, v. 1170, 547–559 (also in arXiv:1001.5283 [astro-ph.SR])

359. Christensen-Dalsgaard J. Helioseismology. *Rev. Mod. Phys.*, 2003, v. 74, 1073–1129 (also in arXiv:astro-ph/0207403).

360. Antia H.M. Solar interior and helioseismology. In, *Lectures on Solar Physics* (H.M. Antia, A. Bhatnagar and R. Ulmschneider, Eds.), Springer, Berlin, 2003, p. 80–126.

361. Bahcall J.N., Pinsonneault M.H. and Basu S. Solar models: Current epoch and time dependences, neutrinos, and helioseismological properties. *Astrophys. J.*, 2001, v. 555, 990–1012.

362. Kosovichev A.G. and Zharkova V.V. X-ray flare sparks quake inside the Sun. *Nature*, 1998, v. 393, 317–318.

363. Fleck B., Brekke P., Haugan S., Duarte L.S., Domingo V., Gurman J.B. and Poland A.I. Four years of SOHO discoveries — Some highlights. *ESA Bulletin*, 2000, v. 102, 68–86.

364. Kuhn J.R., Bush R.I., Schench X., and Scherrer P. The Sun's shape and brightness. *Nature*, 1998, v. 392, no. 6672, 155-157.

365. Charbonneau P., Christensen-Dalsgaard, Henning R., Larsen R.M., Schou J., Thompson M.J. and Tomczyk S. Helioseismic constraints on the structure of the solar tachocline. *Astrophys. J.*, 1999, v. 557, no. 1, 445–460.

366. Antia H.M. and Basu S. Revisiting the solar tachocline: Average properties and temporal variations. *Astrophys. J. Lett.*, 2011, v. 735, no. 2, L45(6 pages).

367. Miesch M.S. Large-scale dynamics of the convection zone and tachocline. *Living Rev. Solar Phys.*, 2005, v. 2, (URL 2, (2005), 1. URL http://www.livingreviews.org/lrsp-2005-1 — Accessed 10/1/2013.

368. Schou J., Antia H.M., Basu S., Bogart R.S., Bush R.I., Chitre S.M., Christensen-Dalsgaard J., di Mauro M.P., Dziembowski W.A., Eff-Darwich A., Gough D.O., Haber D.A., Hoeksema J.T., Howe R., Korzennik S.G, Kosovichev A.G., Larsen R.M., Pijpers F.P., Scherrer P.H., Sekii T., Tarbell T.D., Title A.M., Thompson M.J. and Toomre J.Helioseismic Studies of Differential Rotation in the Solar Envelope by the Solar Oscillations Investigation Using the Michelson Doppler Imager. *Astrophys. J.*, 1998, v. 505, no. 1, 390–417.

369. Nakariakov V.M. and Verwichte E. Coronal waves and oscillations. *Living Rev. Solar Phys.*, 2005, v. 2, 3–65.

370. Roberts B. Progress in coronal seismology. *Proc. IAU: Waves and Oscillations in Solar Atmosphere*, 2007, v. 247, 3–19.

371. De Moortel I. An overview of coronal seismology. *Phil. Trans. R. Soc. A*, 2005, v. 363, 2743–2760.

372. Morton R.J., Verth G., Fedun V., Shelyag S. and Erdélyi R. Evidence for the photospheric excitation of incompressible chromospheric waves. *Astrphys. J.*, 2013, v. 768, no. 1, 17(11 pages).

373. SOHO/LASCO C2 NASA/ESA [2011/05/10 20:00:00 to 2011/05/11 08:00:00 UTC]. These events have been captured in video format and displayed online: e.g. youtube.com/watch?v=igeBrSGk5FA; *Russia Today* youtube.com/watch?NR=1&v=Mat4dWpszoQ&feature=fvwp. (Accessed online on January 29, 2013: Examine beginning at 2011/05/10 2:48:00 to 2011/05/10 4:00:00). NASA provides a tool to generate such films: http://lasco-www.nrl.navy.mil/index.php?p=js_lasco1.

374. Gamow G. The birth and death of the sun: A lucid explanation of stellar evolution and atomic energy. New American Library, New York, N.Y., 1952.

375. Arnett D. Supernovae and nucleosynthesis: An investigation of the history of matter, from the Big Bang to the present. Princeton University Press, Princeton, N.J., 1996.

376. Pagel B. E. J. Nucleosythesis and the chemical evolution of galaxies (2nd Edition), Cambridge University Press, Cambridge, U.K., 2009.

377. Gamow G. Nuclear energy sources and stellar evolution. *Phys. Rev.*, 1938, v. 53, 595–604.

378. Gamow G. Nuclear reactions in stellar evolution. *Nature*, 1939, v. 144, 620–622.

379. Bethe H.A. and Critchfield C.L. The Formation of Deuterons by Proton Combination. *Phys. Rev.*, 1938, v. 54, no. 4, 248–254.

380. Bethe H. A. Energy Production in Stars. *Phys. Rev.*, 1939, v. 55, no. 1, 103.

381. Bethe H.A. Energy Production in Stars. *Phys. Rev.*, 1939, v. 55, no. 5, 434–456.

382. von Weizsäcker C. F. Über Elementumwandlungen in Innern der Sterne II. *Physikalische Zeitschrift*, 1938, v. 39, 633–646.

383. Hoyle F. The synthesis of the elements from hydrogen. *Mon. Not. Roy. Astron. Soc.*, 1946, v. 106, 343–383.

384. Hoyle F. On nuclear reactions occurring in very hot stars. I. The synthesis of elements from carbon to nickel. *Astrophys. J. Suppl. Ser.*, 1954, v. 1, 121–146.

385. Burbidge M., Burbidge G.R., Fowler W.A. and Hoyle F. Synthesis of the elements in stars. *Rev. Mod. Phys.*, 1957, v. 29, no. 4, 547–654.

386. Wallerstein G., Iben I., Parker P., Boesgaard A.M., Hale G.M., Champagne A.E., Barnes C.A., Käppeler F., Smith V.V., Hoffman R.D., Timmes F.X., Sneden C., Boyd R.N., Meyer B.S. and Lambert D.L. Synthesis of the elements in stars: Forty years of progress. *Rev. Mod. Phys.*, 1997, v. 9, no. 4, 995–1084.

387. Bahcall J.N. Neutrinos from the Sun. *Sci. Am.*, 1969, v. 221, no. 1, 28–37.

388. Bahcall J.N. How the Sun shines. http://www.nobelprize.org/nobel_prizes/themes/physics/fusion/.

389. Beers T.C. The first generation stars. *Science*, 2005, v. 309, no. 5733, 390–391.

390. Nickelsen K. and Graβhoff G. Concepts from the Bench: Hans Krebs, Kurt Henseleit and the Urea Cycle. In: "Going Amiss in Experimental Research" (G. Hon, J. Schickore, F. Steinle, Eds.), *Boston Stud. Phil. Science*, 2009, v. 267, 91–117.

391. Krebs H.A., Salvin E. and Johnson A. XX. The formation of citric and α-ketoglutaric acid in the mamalian body. *Biochem. J.*, 1938, v. 32, no. 1, 113-117.

392. Reames D.V. and Ng C.K. Heavy-element abundances in solar energetic particle events. *Astrophys. J.*, 2004, v. 610, 510–522.

393. Kasting J.F. and Grinspoon D.H. The faint young Sun problem: In "The Sun in Time" (C.P. Sonett, M.S. Giampapa, and M.S. Matthews, Eds.), The University of Arizona, Tuscon, 1991, pp. 447–462.

394. Zahnle K., Arndt N., Cockell C., Halliday A., Nisbet E., Selsis F. and Sleep N.H. Emergence of a habitable planet. *Space Sci. Rev.*, 2007, v. 129, 35–78.

395. Solanski S.K. Solar variability and climate change: Is there a link. *Astron. & Geophys.*, 2002, v. 43, no. 5, 5.9–5.13.

396. Lacis A.A., Schmidt G.A., Rind D. and Ruedy R.A. Atmospheric CO_2: Principal control knob governing the Earth's temperature. *Science*, 2010, v. 330, 356–359.

397. Sackmann I.-J. and Boothboyd A. Our Sun V. A bright young Sun consistent with helioseismology and warm temperatures on ancient Earth and Mars. *Astrophys. J.*, 2003, v. 583, 1024–1039.

398. Robitaille P.M. Class-O Wolf-Rayet stars and the birth of the Sun. A new look at stellar evolution. (manuscript in preparation)

399. Murthy V.R. Geochemical evidence for an initially molten Earth. *Phys. Earth Planetary Interiors*, 1992, v. 71, no. 1-2, 46–51.

400. Maunder E.W. Spoerer's Researches on Sun-Spots. *Mon. Not. Roy. Astron. Soc.*, 1890, v. 50, 251–252.

401. Spörer G. Über die Periodicität Sonnenflecken seit dem Jahre 1618, vornehmlich in Bezug auf die heliographische Breite derselben, und Nachweis einer erheblichen Störung dieser Periodicität während eines langen Zeitraumes. la. 4°. *Nova Acta der Ksl. Leopold-Caroline Deutschen Akademie der Natuifoischei*, v. 53, no. 2, 283–324.

402. Spörer. Sur les différences que présentent l'hémisphère nord et l'hémisphère sud du Soleil. *Bull. Astronomique*, 1889, Ser. 1, v. 6, 60–63.

403. Eddy J.A. The Mauder Minimum. *Science*, v. 192, no. 4245, 1189–1202

404. Beckman J.E. and Mahoney T.J. The Mauder minimum and climate change: Have historical records aided current research? *ASP Conf. Ser.*, 1998, v. 153, 212–217.

405. Times of India. Solar Activity Drops to 100 Year Low, Puzzling Scientists. Sept. 18, 2013. http://timesofindia.indiatimes.com/topic/Solar-activity-drops-to-100-year-low.

406. NASA. Big Sun-diving comet discovered. http://solarsystem.nasa.gov/scitech/display.cfm?ST_ID=2504.

407. Haramundanis K. Cecilia Payne-Gaposchkin: An autobiography and other recollections (2nd Edition), Cambridge University Press, Cambridge, U.K., 1996.

408. Rabounski D. and Borissova L. Inside Stars. American Research Press, Rehoboth, NM, 2013. (available online: http://ptep-online.com/index_files/books_files/stars2013.pdf).

n-Valued Refined Neutrosophic Logic and Its Applications to Physics

Florentin Smarandache
University of New Mexico, Math and Sciences Division, 705 Gurley Ave., Gallup, NM 87301, USA.
E-mail:smarand@unm.edu

In this paper we present a short history of logics: from particular cases of 2-symbol or numerical valued logic to the general case of n-symbol or numerical valued logic. We show generalizations of 2-valued Boolean logic to fuzzy logic, also from the Kleene's and Lukasiewicz' 3-symbol valued logics or Belnap's *4*-symbol valued logic to the most general *n-symbol or numerical valued refined neutrosophic logic*. Two classes of neutrosophic norm (*n-norm*) and neutrosophic conorm (*n-conorm*) are defined. Examples of applications of neutrosophic logic to physics are listed in the last section. Similar generalizations can be done for *n-Valued Refined Neutrosophic Set*, and respectively *n-Valued Refined Neutrosopjhic Probability*.

1 Two-Valued Logic

1.1 The Two Symbol-Valued Logic

It is the Chinese philosophy: *Yin and Yang* (or Femininity and Masculinity) as contraries:

Fig. 1: Ying and Yang

It is also the Classical or *Boolean Logic*, which has two symbol-values: truth T and falsity F.

1.2 The Two Numerical-Valued Logic

It is also the Classical or *Boolean Logic*, which has two numerical-values: truth *1* and falsity *0*. More general it is the *Fuzzy Logic*, where the truth (*T*) and the falsity (*F*) can be any numbers in *[0,1]* such that *T + F = 1*.

Even more general, *T* and *F* can be subsets of *[0,1]*.

2 Three-Valued Logic

2.1 The Three Symbol-Valued Logics

1. *Lukasiewicz 's Logic*: True, False, and Possible.
2. *Kleene's Logic*: True, False, Unknown (or Undefined).
3. Chinese philosophy extended to: *Yin*, *Yang*, and *Neuter* (or Femininity, Masculinity, and Neutrality) - as in Neutrosophy. Neutrosophy philosophy was born from neutrality between various philosophies. *Connected with Extenics* (Prof. Cai Wen, 1983), and Paradoxism (F. Smarandache, 1980). *Neutrosophy* is a new branch of philosophy that studies the origin, nature, and scope of neutralities, as well as their interactions with different ideational spectra. This theory considers every notion or idea <A> together with its opposite or negation <antiA> and with their spectrum of neutralities <neutA> in between them (i.e. notions or ideas supporting neither <A> nor <antiA>). The <neutA> and <antiA> ideas together are referred to as <nonA>. Neutrosophy is a generalization of Hegel's dialectics (the last one is based on <A> and <antiA> only). According to this theory every idea <A> tends to be neutralized and balanced by <antiA> and <nonA> ideas - as a state of equilibrium. In a classical way <A>, <neutA>, <antiA> are disjoint two by two. But, since in many cases the borders between notions are vague, imprecise, Sorites, it is possible that <A>, <neutA>, <antiA> (and <nonA> of course) have common parts two by two, or even all three of them as well. *Such contradictions involves Extenics*. Neutrosophy is the base of all neutrosophics and it is used in engineering applications (especially for software and information fusion), medicine, military, airspace, cybernetics, physics.

2.2 The Three Numerical-Valued Logic

1. *Kleene's Logic*: True (*1*), False (*0*), Unknown (or Undefined) (*1/2*), and uses "min" for $\wedge$, "max" for $\vee$, and "1-" for negation.
2. More general is the *Neutrosophic Logic* [Smarandache, 1995], where the truth (*T*) and the falsity (*F*) and the indeterminacy (*I*) can be any numbers in *[0, 1]*, then $0 \leq T + I + F \leq 3$. More general: Truth (*T*), Falsity (*F*), and Indeterminacy (*I*) are standard or nonstandard subsets of the nonstandard interval $]^{-}0, 1^{+}[$.

3 Four-Valued Logic

3.1 The Four Symbol-Valued Logic

1. It is *Belnap's Logic*: True (*T*), False (*F*), Unknown (*U*), and Contradiction (*C*), where *T*, *F*, *U*, *C* are symbols,

not numbers. Below is the Belnap's conjunction operator table:

∩	F	U	C	T
F	F	F	F	F
U	F	U	F	U
C	F	F	C	C
T	F	U	C	T

Restricted to *T,F,U*, and to *T,F,C*, the Belnap connectives coincide with the connectives in Kleene's logic.

2. Let *G* = Ignorance. We can also propose the following two 4-Symbol Valued Logics: *(T, F, U, G)*, and *(T, F, C, G)*.
3. *Absolute-Relative 2-, 3-, 4-, 5-, or 6-Symbol Valued Logics* [Smarandache, 1995]. Let T_A be truth in all possible worlds (according to Leibniz's definition); T_R be truth in at last one world but not in all worlds; and similarly let I_A be indeterminacy in all possible worlds; I_R be indeterminacy in at last one world but not in all worlds; also let F_A be falsity in all possible worlds; F_R be falsity in at last one world but not in all worlds; Then we can form several Absolute-Relative 2-, 3-, 4-, 5-, or 6-Symbol Valued Logics just taking combinations of the symbols T_A, T_R, I_A, I_R, F_A, and F_R. As particular cases, very interesting would be to study the Absolute-Relative 4-Symbol Valued Logic (T_A, T_R, F_A, F_R), as well as the Absolute-Relative 6-Symbol Valued Logic (T_A, T_R, I_A, I_R, F_A, F_R).

3.2 Four Numerical-Valued Neutrosophic Logic

Indeterminacy I is refined (split) as U = Unknown, and C = contradiction. T, F, U, C are subsets of [0, 1], instead of symbols; This logic generalizes Belnap's logic since one gets a degree of truth, a degree of falsity, a degree of unknown, and a degree of contradiction. Since $C = T \wedge F$, this logic involves the Extenics.

4 Five-Valued Logic

1. Five Symbol-Valued Neutrosophic Logic [Smarandache, 1995]: Indeterminacy I is refined (split) as U = Unknown, C = contradiction, and G = ignorance; where the symbols represent:
   T = truth;
   F = falsity;
   U = neither T nor F (undefined);
   $C = T \wedge F$, which involves the Extenics;
   $G = T \vee F$.
2. If *T, F, U, C,* G are subsets of *[0, 1]* then we get:a *Five Numerical-Valued Neutrosophic Logic*.

5 Seven-Valued Logic

1. *Seven Symbol-Valued Neutrosophic Logic* [Smarandache, 1995]:
   I is refined (split) as *U, C, G*, but *T* also is refined as T_A = absolute truth and T_R = relative truth, and *F* is refined as F_A = absolute falsity and F_R = relative falsity. Where: U = neither (T_A or T_R) nor (F_A or F_R) (i.e. undefined); $C = (T_A or T_R) \wedge (F_A or F_R)$ (i.e. Contradiction), which involves the Extenics;
   $G = (T_A or T_R) \vee (F_A or F_R)$ (i.e. Ignorance). All are symbols.
2. But if T_A, T_R, F_A, F_R, *U, C, G* are subsets of *[0, 1]*, then we get a *Seven Numerical-Valued Neutrosophic Logic*.

6 n-Valued Logic

1. *The n-Symbol-Valued Refined Neutrosophic Logic* [Smarandache, 1995]. In general:
   T can be split into many types of truths: $T_1, T_2, ..., T_p$, and *I* into many types of indeterminacies: $I_1, I_2, ..., I_r$, and *F* into many types of falsities: $F_1, F_2, ..., F_s$, where all $p, r, s \geq 1$ are integers, and $p + r + s = n$.
   All subcomponents T_j, I_k, F_l are symbols for $j \in \{1, 2..., p\}$, $k \in \{1, 2..., r\}$, and $l \in \{1, 2..., s\}$.
   If at least one $I_k = T_j \wedge F_l$ =contradiction, we get again the Extenics.
2. *The n-Numerical-Valued Refined Neutrosophic Logic.* In the same way, but all subcomponents T_j, I_k, F_l are not symbols, but subsets of *[0,1]*, for all $j \in \{1, 2..., p\}$, all $k \in \{1, 2..., r\}$, and all $l \in \{1, 2..., s\}$. If all sources of information that separately provide neutrosophic values for a specific subcomponent are independent sources, then in the general case we consider that each of the subcomponents T_j, I_k, F_l is independent with respect to the others and it is in the non-standard set $]^-0, 1^+[$. Therefore per total we have for crisp neutrosophic value subcomponents T_j, I_k, F_l that:

$$^-0 \leq \sum_{j=1}^{p} T_j + \sum_{k=1}^{r} I_k + \sum_{l=1}^{s} F_l \leq n^+ \qquad (1)$$

   where of course $n = p + r + s$ as above. If there are some dependent sources (or respectively some dependent subcomponents), we can treat those dependent subcomponents together. For example, if T_2 and I_3 are dependent, we put them together as $^-0 \leq T_2 + I_3 \leq 1^+$. The non-standard unit interval $]^-0, 1^+[$, used to make a distinction between absolute and relative truth/ indeterminacy /falsehood in philosophical applications, is replace for simplicity with the standard (classical) unit interval $[0, 1]$ for technical applications.
   For at least one $I_k = T_j \wedge F_l$ = contradiction, we get again the Extenics.

7 n-Valued Neutrosophic Logic Connectors

1. **n-Norm and n-Conorm defined on combinations of t-Norm and t-Conorm**
   The n-norm is actually the neutrosophic conjunction operator, NEUTROSOPHIC AND ($\wedge_n$); while the n-conorm is the neutrosophic disjunction operator, NEUTROSOPHIC OR ($\vee_n$).
   One can use the t-norm and t-conorm operators from the fuzzy logic in order to define the **n-norm** and respectively **n-conorm** in neutrosophic logic:

$$\begin{aligned} &n-norm((T_j)_{j=\{1,2,...,p\}},\\ &(I_k)_{k=\{1,2,...,r\}},(F_l)_{l=\{1,2,...,s\}})\\ &=([t-norm(T_j)]_{j=\{1,2,...,p\}},\\ &\quad [t-conorm(I_k)]_{k=\{1,2,...,r\}},\\ &\quad [t-conorm(F_l)]_{l=\{1,2,...,s\}}) \end{aligned} \tag{2}$$

   and

$$\begin{aligned} &n-conorm((T_j)_{j=\{1,2,...,p\}},(I_k)_{k=\{1,2,...,r\}},\\ &(F_l)_{l=\{1,2,...,s\}})\\ &=([t-conorm(T_j)]_{j=\{1,2,...,p\}},\\ &\quad [t-norm(I_k)]_{k=1,2,...,r},\\ &\quad [t-norm(F_l)]_{l=1,2,...,s}) \end{aligned} \tag{3}$$

   and then one normalizes if needed.
   Since the n-norms/n-conorms, alike t-norms/t-conorms, can only approximate the inter-connectivity between two n-Valued Neutrosophic Propositions, there are many versions of these approximations.
   For example, for the n-norm: the indeterminate (sub)components I_k alone can be combined with the t-conorm in a pessimistic way [i.e. lower bound], or with the t-norm in an optimistic way [upper bound]; while for the n-conorm: the indeterminate (sub)components I_k alone can be combined with the t-norm in a pessimistic way [i.e. lower bound], or with the t-conorm in an optimistic way [upper bound].
   In general, if one uses in defining an n-norm/n-conorm for example the t-norm $min\{x, y\}$ then it is indicated that the corresponding t-conorm used be $max\{x, y\}$; or if the t-norm used is the product $x\dot{y}$ then the corresponding t-conorm should be $x+y-x\dot{y}$; and similarly if the t-norm used is $max\{0,\ x+y-1\}$ then the corresponding t-conorm should be $min\{x+y,\ 1\}$; and so on.
   Yet, it is still possible to define the n-norm and n-conorm using different types of t-norms and t-conorms.
2. **N-norm and n-conorm based on priorities**
   For the *n-norm* we can consider the priority: T<I<F, where the subcomponents are supposed to conform with similar priorities, i.e.

$$\begin{aligned} &T_1<T_2<...<T_p<I_1<I_2<...\\ &<I_r<F_1<F_2<...<F_s. \end{aligned} \tag{4}$$

   While for the **n-conorm** one has the opposite priorities: T>I>F, or for the refined case:

$$\begin{aligned} &T_1>T_2>...>T_p>I_1>I_2>...\\ &>I_r>F_1>F_2>...>F_s. \end{aligned} \tag{5}$$

   By definition A<B means that all products between A and B go to B (the bigger).
   Let's say, one has two neutrosophic values in simple (non-refined case):

$$(T_x, I_x, F_x) \tag{6}$$

   and

$$(T_y, I_y, F_y) \tag{7}$$

   Applying the n-norm to both of them, with priorities T < I < F, we get:

$$\begin{aligned} &(T_x, I_x, F_x)\wedge_n(T_y, I_y, F_y)\\ &=(T_xT_y, T_xI_y+T_yI_x+I_xI_y,\\ &\quad T_xF_y+T_yF_x+I_xF_y+I_yF_x+F_xF_y). \end{aligned} \tag{8}$$

   Applying the n-conorm to both of them, with priorities T > I > F, we get:

$$\begin{aligned} &(T_x, I_x, F_x)\vee_n(T_y, I_y, F_y)\\ &=(T_xT_y+T_xI_y+T_yI_x+T_xF_y+T_yF_x,\\ &\quad I_xI_y+I_xF_y+I_yF_x, F_xF_y). \end{aligned} \tag{9}$$

   In a lower bound (pessimistic) n-norm one considers the priorities T < I < F, while in an upper bound (optimistic) n-norm one considers the priorities I < T < F.
   Whereas, in an upper bound (optimistic) n-conorm one considers T > I > F, while in a lower bound (pessimistic) n-conorm one considers the priorities T>F>I.
   Various priorities can be employed by other researchers depending on each particular application.

8 Particular Cases

If in *6 a)* and *b)* one has all $I_k = 0, k = \{1, 2, ..., r\}$, we get the **n-Valued Refined Fuzzy Logic**.

If in *6 a)* and *b)* one has only one type of indeterminacy, i.e. k=1, hence $I_1 = I > 0$, we get the **n-Valued Refined Intuitionistic Fuzzy Logic**.

9 Distinction between Neutrosophic Physics and Paradoxist Physics

Firstly, we make a distinction between Neutrosophic Physics and Paradoxist Physics.

1. **Neutrosophic Physics**
   Let <A> be a physical entity (i.e. concept, notion, object, space, field, idea, law, property, state, attribute, theorem, theory, etc.), <antiA> be the opposite of <A>, and <neutA> be their neutral (i.e. neither <A> nor <antiA>, but in between).

Neutrosophic Physics is a mixture of two or three of these entities <A>, <antiA>, and <neutA> that hold together.

Therefore, we can have neutrosophic fields, and neutrosophic objects, neutrosophic states, etc.

2. **Paradoxist Physics**
Neutrosophic Physics is an extension of Paradoxist Physics, since Paradoxist Physics is a combination of physical contradictories <A> and <antiA> only that hold together, without referring to their neutrality <neutA>. Paradoxist Physics describes collections of objects or states that are individually characterized by contradictory properties, or are characterized neither by a property nor by the opposite of that property, or are composed of contradictory sub-elements. Such objects or states are called paradoxist entities.

These domains of research were set up in the *1995* within the frame of neutrosophy, neutrosophic logic/ set/probability/statistics.

10 n-Valued Refined Neutrosophic Logic Applied to Physics

There are many cases in the scientific (and also in humanistic) fields that two or three of these items <A>, <antiA>, and <neutA> simultaneously coexist.

Several **Examples** of paradoxist and neutrosophic entities:

- anions in two spatial dimensions are arbitrary spin particles that are neither bosons (integer spin) nor fermions (half integer spin);
- among possible Dark Matter candidates there may be exotic particles that are neither Dirac nor Majorana fermions;
- mercury (Hg) is a state that is neither liquid nor solid under normal conditions at room temperature;
- non-magnetic materials are neither ferromagnetic nor anti-ferromagnetic;
- quark gluon plasma (QGP) is a phase formed by quasi-free quarks and gluons that behaves neither like a conventional plasma nor as an ordinary liquid;
- unmatter, which is formed by matter and antimatter that bind together (F. Smarandache, 2004);
- neutral kaon, which is a pion and anti-pion composite (R. M. Santilli, 1978) and thus a form of unmatter;
- neutrosophic methods in General Relativity (D. Rabounski, F. Smarandache, L. Borissova, 2005);
- neutrosophic cosmological model (D. Rabounski, L. Borissova, 2011);
- neutrosophic gravitation (D. Rabounski);
- qubit and generally quantum superposition of states;
- semiconductors are neither conductors nor isolators;
- semi-transparent optical components are neither opaque nor perfectly transparent to light;
- quantum states are metastable (neither perfectly stable, nor unstable);
- neutrino-photon doublet (E. Goldfain);
- the "multiplet" of elementary particles is a kind of "neutrosophic field" with two or more values (E. Goldfain, 2011);
- A "neutrosophic field" can be generalized to that of operators whose action is selective. The effect of the neutrosophic field is somehow equivalent with the "tunneling" from the solid physics, or with the "spontaneous symmetry breaking" (SSB) where there is an internal symmetry which is broken by a particular selection of the vacuum state (E. Goldfain). Etc.

Many types of logics have been presented above. For the most general logic, the n-valued refined neutrosophic logic, we presented two classes of neutrosophic operators to be used in combinations of neutrosophic valued propositions in physics.

Similar generalizations are done for **n-Valued Refined Neutrosophic Set, and respectively n-Valued Refined Neutrosophic Probability**.

Submitted on October 24, 2013 / Accepted on October 26, 2013

References

1. Dubois D. Uncertainty Theories, Degrees of Truth and Epistemic States, http://www.icaart.org/Documents/Previous-Invited-Speakers/2011/ICAART2011-Dubois.pdf
2. Smarandache F. (Editor). Proceedings of the Introduction to Neutrosophic Physics: Unmatter and Unparticle — International Conference, Zip Publ., Columbus, 2011.
3. Rabounski D., Smarandache F., Borisova L. Neutrosophic Methods in General Relativity. Neutrosophic Book Series, 10. *Hexis, Phoenix*, AZ, 2005. (Re-printed in Russian as: Netrosofskie Metody v Obshchey Teorii Otnositelnosti. *Hexis, Phoenix*, AZ, 2005.)
4. Smarandache F. Neutrosophic Logic and Set, mss., http://fs.gallup.unm.edu/neutrosophy.htm, 1995.
5. Smarandache F. A Unifying Field in Logics: Neutrosophic Field. *Multiple-Valued Logic / An International Journal*, 2002, v.8, no.3, 385–438. (This issue of the journal is dedicated to Neutrosophy and Neutrosophic Logic.)
6. Rivieccio U. Neutrosophic logics: Prospects and problems. *Fuzzy Sets and Systems*, 2008, v. 159, issue 14, 1860–1868.
7. Smarandache F. An Introduction to the Neutrosophic Probability Applied in Quantum Statistics. *Bulletin of Pure and Applied Sciences, Physics* 2003, v. 22D, no. 1, 13–25.
8. Smarandache F. Neutrosophic Set-A Generalization of the Intuitionistic Fuzzy Set. *Intern. Journal of Pure and Applied Mathematics*, 2005, v. 24, no. 3, 287–297.
9. Dezert J. Open questions on neutrosophic inference. Neutrosophy and neutrosophic logic. *Multiple-Valued Logic / An International Journal*, 2002, v. 8, no. 3, 439–472.
10. Webster's Online Dictionary, Paraconsistent probability (neutrosophic probability). http://www.websters-online-dictionary.org

Path Distribution Energy and Possible Consequences

Janez Špringer

Cankarjeva Cesta 2, 9250 Gornja Radgona, Slovenia, EU. E-mail: info@lekarna-springer.si

Previously (*Progress in Physics, 2013, v. 4, 83–84*) one investigated the geometric distribution of the frequencies of the path of the electron in the ground state of Hydrogen atom. In this paper one shows that the resulting difference detected on the fifth decimal of the inverse fine structure constant is accompanied by the difference in the quantized energy up to 0.04 eV. The difference in charge as well as energy of the distributed and non-distributed electrons could explain the origin of van der Waals intermolecular interactions.

1 Theoretical background

The distribution of the path of the electron changes the inverse fine structure constant [1]. Let us see what is accompanied to that change. The inverse fine structure constant can be expressed as:

$$\alpha^{-1} = \frac{2\varepsilon_0 hc}{e^2}\,. \tag{1}$$

The energy equivalent of the mass of the electron E_e can be expressed as [2]:

$$E_e = m_e c^2 = \frac{e^2}{4\pi\varepsilon_0 r_e}\,. \tag{2}$$

The inverse fine structure constant α^{-1} and the energy equivalent of the mass of the electron E_e are in inverse proportion since combining (1) and (2) the next relation is given:

$$\alpha^{-1} = \frac{hc}{2\pi E_e r_e}\,. \tag{3}$$

Other parameters staying untouched the inverse fine structure constant α^{-1} is changed due to the change of the electron charge e and consequently the energy equivalent of the mass E_e. Energetically more favorable is the greater inverse fine structure constant α^{-1} since it belongs to the smaller charge e and energy equivalent E_e. Therefore the proposed distributed path of the electron in the ground state of Hydrogen atom [1] is more favorable than non-distributed one. Having greater α^{-1} possesses lower E_e. The most favorable is the infinite-sided distribution with the largest α^{-1} and the lowest E_e. Energies of the discrete distributions are quantized. The difference in energy between the non-distributed electron E_0 and on the arbitrary number of the even-sides k distributed electron E_k is given by:

$$\Delta E_k = E_0 - E_k. \tag{4}$$

Because of the inverse proportion of α^{-1} and E_e holds:

$$\frac{\alpha^{-1}_{k-sided}}{\alpha^{-1}_{0-sided}} = \frac{E_0}{E_k}\,. \tag{5}$$

The difference in energy is then expressed as:

$$\Delta E_k = \left(\frac{\alpha^{-1}_{k-sided}}{\alpha^{-1}_{0-sided}} - 1\right) E_0\,. \tag{6}$$

The difference in energy between the energy equivalents of the mass of the electron at the different number of sides of distribution ΔE_k (4) is also the difference of the distribution energies ΔE_d:

$$\Delta E_k = \Delta E_d = \Delta E_{0-distribution} - E_{k-distribution}\,. \tag{7}$$

The distribution energy of the non-distribution is assumed to be zero:

$$E_{0-distribution} = 0. \tag{8}$$

So the distribution energy of the path of the electron of the arbitrary *k-sided* distribution is given by:

$$E_{k-distribution} = -\Delta E_d = -\Delta E_k. \tag{9}$$

The negative distribution energy means that energy is released in the case when the electron path becomes distributed, and on the contrary, the energy is spent in the case when the electron path becomes non-distributed. The distribution of the path of the electron does not need to be atom-radius dependent (it is distribution-radius dependent) [1] so what applies for Hydrogen atom could hold true also for other atoms.

2 Calculation of the Distribution Energy

The non-distribution energy $E_{0-distribution}$ is zero by definition (8).

On the two decimals rounded energy of the two-sided distribution can be calculated with the help of equations (6) and (9) knowing the CODATA value of the energy equivalent of the mass of the electron $E_e = 510998.91$ eV, and the appropriate distributed inverse fine structure constants $\alpha^{-1}_{0-sided} = 137.036006$ and $\alpha^{-1}_{2-sided} = 137.036014$ [1]:

$$E_{2-distribution} = -0.03 \text{ eV}. \tag{10}$$

On the two decimals rounded energy of the infinite-sided distribution can be calculated in the same way knowing the inverse fine structure constant $\alpha^{-1}_{\infty-sided} = 137.036018$ [1]:

$$E_{\infty-distribution} = -0.04 \text{ eV}. \quad (11)$$

The infinite number of the quantized distribution energies in the range of 0.04 eV can be calculated on all *k*-sides of the ground state of Hydrogen atom. Of course this paper brings no statement of how many of them are physically true.

3 Instead of conclusion

The proposed quantized distribution energies of the electron seem to have physico-chemical consequences. Ranged up to 0.04 eV (10), (11) are of the same order of magnitude as the typical energies from 0.4 kJ/mol to 4 kJ/mol of the van der Waals interaction between atoms [3]. Indeed:

$$\frac{4 \text{ kJ}}{\text{mol}} \approx \frac{0.04 \text{ eV}}{\text{molecule}}. \quad (12)$$

The different energy and charge of the distributed and non-distributed electrons could explain the origin of the mentioned intermolecular interactions.

Dedication

This fragment is dedicated to my wife Ivka.

Submitted on October 7, 2013 / Accepted on October 23, 2013

References

1. Špringer J. Geometric Distribution of Path and Fine Structure. *Progress in Physics*, 2013, v. 4, 83–84.
2. Haken H., Wolf H.C., Brewer W.D. The Physics of Atoms and Quanta: Introduction to Experiments and Theory. 2005, Springer. p. 70.
3. http://chemwiki.ucdavis.edu/Physical_Chemistry/Quantum_Mechanics/Atomic_Theory/Intermolecular_Forces/Van_Der_Waals_Interactions. Retrieved October 2013.

LETTERS TO PROGRESS IN PHYSICS

Open Letter by the Editor-in-Chief:
Declaration of Academic Freedom (Scientific Human Rights)
The Slovene Translation

Deklaracija akademske svobode
(Človekove pravice na znanstvenem področju)

Original text published in English: *Progress in Physics*, 2006, v.1, 57–60. Izvirno angleško besedilo: Dmitri Rabounski, glavni urednik revije *Progress in Physics*. E-mail: rabounski@yahoo.com. V slovenščino prevedla Janez Špringer in Darinka Špringer (Slovenija, EU), E-mail: info@lekarna-springer.si.

1 člen: Preambula

Začetek 21. stoletja bolj kot katero koli drugo obdobje v zgodovini človeštva odseva globino in pomembnost vloge znanosti in tehnologije pri stvareh, ki nas kot ljudi zadevajo.

Nadvse prevladujoča narava moderne znanosti in tehnologije je privedla do splošnega prepričanja, da je bodoča pomembnejša odkritja mogoče doseči v glavnem ali zgolj v velikih vladnih ali korporativno financiranih raziskovalnih skupinah, ki imajo na voljo nezaslišano drag instrumentarij in nebroj pomožnega osebja.

Običajna predstava pa je vendarle izmišljena in lažno zrcali dejansko naravo, kako se do znanstvenih odkritij v resnici pride. Veliki in dragi tehnološki projekti, kakor koli že so zapleteni, niso nič drugega kot izid uporabe poglobljenih znanstvenih uvidov manjših skupin predanih raziskovalcev ali samostojnih znanstvenikov, ki pogosto delajo v odmaknjenosti. Znanstvenik, ki dela sam, je sedaj in bo v prihodnje, kakor je bil že v preteklosti, sposoben priti do odkritja, ki lahko bistveno vpliva na usodo človeštva in spremeni obličje celotnega planeta, kjer tako nepomembno prebivamo.

Do velikih odkritij se po navadi dokopljejo posamezniki, ki delajo na podrejenih delovnih mestih znotraj vladnih agencij, raziskovalnih in izobraževalnih ustanov ali komercialnih podjetij. Posledično direktorji podjetij in institucij raziskovalca pogosto omejujejo in zatirajo, saj stremijo k drugim ciljem in želijo znanstveno raziskavo nadzorovati, odkritje pa uporabiti organizaciji ali sebi v prid ter sami sebe poveličati.

Zgodovina znanstvenih odkritij je prepolna zatiranja in posmehovanja, ki ju je izvajala sprevrnjena elita; šele v poznejših letih so bili primeri razkriti v pravi luči zaradi nezadržnega pohoda praktične nujnosti in intelektualnega razsvetljenja. Takisto je znanost omadeževana in oskrunjena s plagiatorstvom in namernim popačenjem, ki so ju zaradi zavisti in pohlepa izvajali brezobzirneži. In tako je tudi danes.

Namen te deklaracije je ohraniti in nadaljevati temeljno doktrino, da mora znanstveno raziskovanje potekati tako brez prikritega kot odkritega represivnega vpliva birokratskih, političnih, religioznih in kapitalskih smernic in da znanstveno ustvarjanje ni nič manjša človekova pravica kot druge takšne pravice in silni upi, zapisani v mednarodnih sporazumih in mednarodnem pravu.

Vsi znanstveniki naj spoštujejo to deklaracijo v znak solidarnosti z mednarodno znanstveno skupnostjo in z namenom, da bi se prebivalcem sveta omogočile pravice za neovirano znanstveno ustvarjanje na podlagi individualnih sposobnosti in naravnanosti. Za napredek v znanosti gre, za to naj si kot spodobni državljani prizadevajo po svojih najboljših močeh v tem nespodobnem svetu, in za blagor človeštva. Znanost in tehnologija sta bili že predolgo žrtvi zatiranja.

2 člen: Kdo je znanstvenik

Znanstvenik je oseba, ki se ukvarja z znanostjo. Vsakdo, ki sodeluje z znanstvenikom pri razvijanju in predlaganju idej in podatkov pri raziskavi ali njeni uporabi, je tudi znanstvenik. Formalna izobrazba ni predpogoj za to, da kdo postane znanstvenik.

3 člen: Kje nastaja znanost

Znanstveno raziskavo je mogoče izvajati na sploh kjerkoli, denimo v službi, med potekom formalnega izobraževanja in med sponzoriranim akademskim programom, tako v skupinah ali kot posameznik, ki neodvisno raziskuje doma.

4 člen: Svobodna izbira raziskovalne teme

Mnogim znanstvenikom, ki se potegujejo za višje strokovne nazive ali so udeleženi pri drugih raziskovalnih programih v akademskih ustanovah, kot so na primer univerze in šole za izpopolnjevalni študij, starejši akademiki in/ali administratorji preprečujejo delo na raziskovalni temi po lastni izbiri. Ne sicer zaradi primanjkljaja ustrezne opreme in prostorov, pač pa iz razloga, da akademska hierarhija in/ali drugi uradniki enostavno ne odobravajo takšnih raziskav, saj bi lahko prevrnile prevladujočo dogmo in favorizirane teorije ali celo ogrozile financiranje drugih projektov, ki jim predlagana

raziskava nemara odvzame veljavo. Avtoriteta ortodoksne večine pogosto onemogoči moteč raziskovalni projekt, tako da niti avtoriteta niti proračun nista prizadeta. Ta vsakdanja praksa je namerna ovira za svobodno znanstveno misel. Ker je povsem neznanstvena in zločinska, ji ne moremo gledati skozi prste.

Znanstvenik, ki dela za katero koli akademsko ustanovo, avtoriteto ali agencijo, mora biti pri izbiri raziskovalne teme popolnoma svoboden, omejen sme biti le z materialno podporo in intelektualnimi sposobnostmi, ki jih zmore nuditi izobraževalna ustanova, agencija ali avtoriteta. Če znanstvenik raziskuje kot član raziskovalne skupine, naj bodo vloge direktorjev raziskav in vodij skupin le svetovalne in posvetovalne narave glede na izbiro ustrezne raziskovalne teme znanstvenikov v skupini.

5 člen: Svobodna izbira raziskovalnih metod

Pri izvedbi raziskovalnega programa znotraj akademskega okolja administrativno osebje ali starejši akademiki pogosto silijo k uporabi drugačnih raziskovalnih metod od tistih, ki jih je znanstvenik sam izbral. Razlogov za to ni mogoče poiskati drugje kot v osebnih preferencah, pristranskosti, institucionalni politiki, uredniških zapovedih ali kolektivni avtoriteti. Takšna precej razširjena praksa je namerno zanikanje miselne svobode in ni dopustna.

Nekomercialni ali akademski znanstvenik ima pravico obdelati raziskovalno temo na kateri koli razumen način in s kakršnimi koli razumnimi sredstvi, za katera sam meni, da so najučinkovitejša. Končna odločitev o načinu poteka raziskave je le znanstvenikova.

Če nekomercialni ali akademski znanstvenik deluje kot član nekomercialne ali akademske skupine znanstvenikov, naj imajo vodje projektov in direktorji raziskav zgolj svetovalne in posvetovalne pravice in naj ne slabijo, omejujejo in na kakršen koli drug način posegajo v uporabo raziskovalne metode in obdelavo raziskovalne teme znanstvenika znotraj skupine.

6 člen: Svobodna udeležba in sodelovanje pri raziskovanju

Prakso moderne znanosti bremeni značilno institucionalno rivalstvo, ki ga spremljata osebna zavist in ohranjanje ugleda za vsako ceno brez upoštevanja raziskovalne resničnosti. To dejstvo znanstvenikom pogosto preprečuje sodelovanje s kompetentnimi kolegi, tako nameščenimi v rivalskih ustanovah kot drugimi brez sleherne akademske pripadnosti. Tudi takšna praksa je namerna ovira znanstvenemu napredku.

V primeru, da nekomercialni znanstvenik potrebuje pomoč drugega znanstvenika in slednji vanjo privoli, se ga sme brez zadržka prositi za kakršno koli in vsakršno pomoč pod pogojem, da nudenje pomoči ne presega okvira raziskovalnega proračuna. Če pomoč ni vezana na proračun, se sme znanstvenik svobodno odločiti zanjo in pritegniti k sodelovanju pomočnika povsem po lastni presoji brez kakršnega koli vmešavanja kogar koli.

7 člen: Svobodno nestrinjanje pri znanstveni razpravi

Zaradi skrivnega ljubosumja in pridobitniškega interesa moderna znanost prezira odprto razpravo in odločno preganja tiste znanstvenike, ki dvomijo o ortodoksnih stališčih. Znanstveniki z izrednimi sposobnostmi, ki opozorijo na pomanjkljivosti v trenutni teoriji ali interpretaciji podatkov, so zelo pogosto označeni za čudake, saj je tako mogoče njihova stališča z lahkoto ignorirati. Javno in zasebno so zasmehovani, sistematično pa se jim onemogoča tudi udeležba na znanstvenih kongresih, seminarjih in kolokvijih, tako da njihove ideje ostanejo brez poti do občinstva. Načrtno ponarejanje podatkov in napačno interpretiranje teorij današnje dni brezobzirnežem pogosto služita kot orodje za prikrivanje tako tehničnih kot znanstvenih dejstev. Izoblikovali so se mednarodni odbori znanstvenih nastopačev, ki prirejajo in usmerjajo mednarodne kongrese, kjer smejo svoje referate ne glede na vsebinsko kakovost predstavljati le njihovi privrženci. Ti odbori z zatekanjem k prevaram in lažem iz javne blagajne izvlečejo velikanske vsote denarja za financiranje svojih sponzoriranih projektov. Da se denar lahko še naprej nemoteno steka na račune za njihove projekte in jim tako zagotavlja dobro plačane službe, se vsakršno znanstveno utemeljeno nasprotovanje njihovim predlogom utiša z vsemi njim razpoložljivimi sredstvi. Oporečnim znanstvenikom se na podlagi njihovih ukazov vročajo odpovedi; drugim se s pomočjo mreže skorumpiranih pajdašev prepreči dostop do akademskih imenovanj. V spet drugih okoliščinah se onemogočijo kandidature pri programih za pridobitev višje stopnje strokovnosti, na primer doktorskega naziva, in to zaradi izražanja idej, ki spodkopavajo moderno teorijo, ne glede na to, za kakšno staro ortodoksno teorijo že gre. Temeljno dejstvo, ki pravi, da nobena znanstvena teorija ni dokončna in nedotakljiva, in je zategadelj odprta za razpravo in ponovno preverbo, popolnoma ignorirajo. Prav tako ignorirajo dejstvo, da ima nek pojav več mogočih razlag, in se škodoželjno obregnejo ob vsako, ki ni v skladu z ortodoksnim mnenjem; da pa bi opravičili svoja pristranska mnenja, se brez obotavljanja poslužujejo neznanstvene argumentacije.

Vsi znanstveniki naj imajo pravico do svobodne razprave o svojih raziskavah in raziskavah drugih. Naj bodo brez strahu pred javnim ali zasebnim objektivno neutemeljenim posmehom oziroma brez bojazni, da bodo na podlagi neupravičenih navedb postali tarče obtoževanja, omalovaževanja, poniževanja in siceršnjega zaničevanja. Nihče naj ne bo postavljen v položaj, kjer bi bila zaradi izražanja znanstvenih stališč ogrožena njegovo preživljanje in ugled. Svoboda znanstvenega izražanja naj bo najpomembnejša. Uporaba avtoritete za ovržbo znanstvenih dokazov ni znanstvena in naj se je ne uporablja za zavezovanje ust, zatiranje, ustrahovanje, preganjanje ali kakršno koli drugo

priganjanje oziroma ustvarjanje pritiska na znanstvenika. Namerno zamolčanje znanstvenih dejstev ali dokazov z dejanjem ali opustitvijo dejanja in namerno prirejanje podatkov v podporo dokazom ali za diskreditiranje nasprotnega stališča je znanstvena prevara, ki velja za znanstveni zločin. Načelo dokazov naj vodi vso znanstveno razpravo, najsi bodo ti dokazi praktični, teoretični ali preplet obojega.

8 člen: Svobodno objavljanje znanstvenih dognanj

Obžalovanja vredna cenzura znanstvenih člankov je postala današnje dni stalna praksa uredniških odborov pomembnejših revij in elektronskih arhivov ter navez njihovih domnevnih strokovnih razsodnikov. Razsodnike zvečine varuje anonimnost, tako da avtor ne more preveriti njihove domnevne strokovnosti. Objava znanstvenega dela se današnje dni rutinsko zavrne v primeru, ko se avtor ne strinja s preferenčno teorijo in večinsko pravovernostjo ali jima celo nasprotuje. Brez vsebinskih razlogov se marsikateri članek samodejno zavrne, zgolj če je njegov avtor na seznamu nečislanih pri urednikih, razsodnikih ali drugih strokovnih cenzorjih. Obstaja črni seznam disidentskih znanstvenikov, s katerega vsebino so seznanjeni povezani uredniški odbori. Vse to je velikanska pristranskost in graje vredno zatiranje svobodnega mišljenja ter si zasluži obsodbo mednarodne znanstvene skupnosti.

Vsi znanstveniki naj imajo pravico predstaviti dognanja svojih znanstvenih raziskav ali v celoti ali delno na ustreznih znanstvenih konferencah ter jih objaviti v tiskanih znanstvenih revijah, elektronskih arhivih in drugih medijih. Nobenemu znanstveniku naj se ne zavrne objava članka ali poročila, predloženega za objavo v znanstveni reviji, elektronskem arhivu ali drugem mediju zgolj zato, ker njegovo delo zaseje dvom o trenutnem večinskem prepričanju, je v nasprotju s pogledi uredniškega odbora, spodkopava temelje trenutnih ali bodočih raziskovalnih projektov drugih znanstvenikov ali je v nasprotju s kakršno koli politično dogmo, verskim prepričanjem in osebnim mnenjem drugega. Prav tako naj ne bo noben znanstvenik uvrščen na črno listo ali kako drugače cenzuriran, nihče pa mu tudi naj ne preprečuje objavljanja.

Noben znanstvenik naj zaradi obljube prejemanja daril ali kakršnih koli podkupnin ne ovira, spreminja ali se kako drugače vpleta v objavljanje del drugega znanstvenika.

9 člen: Soavtorstvo znanstvenih del

V znanstvenih krogih je komaj še skrito dejstvo, da ima veliko soavtorjev raziskovalnih člankov malo ali skoraj nič opraviti z objavljeno raziskavo. Veliko nadzornikov podiplomskih študentov, denimo, se ne brani pripisa za soavtorstvo člankov, ki so jih pod njihovim formalnim nadzorom napisali podrejeni znanstveniki. V veliko takšnih primerih je dejanski pisec inteligentnejši od formalnega nadzornika. V drugih primerih, spet zaradi slave, slovesa, denarja, ugleda ali česar podobnega, si tretje osebe pripisujejo soavtorstvo člankov. Pravim avtorjem takšnih člankov preostane le ugovor, z njim pa tvegajo, da bodo na nek način kaznovani, celo v obliki zavrnitve pri kandidiranju za višji raziskovalni naziv ali sodelovanje v raziskovalni skupini, kot se pogosto dogaja. Veliko jih je bilo v takšnih okoliščinah v resnici zavrnjenih. Te pretresljive prakse ne moremo več dopustiti. Avtorstvo naj se pripiše le za raziskavo odgovornim osebam.

Noben znanstvenik naj ne predlaga drugemu, ki ni sodeloval pri raziskavi, da bi postal soavtor članka, in noben znanstvenik naj ne dovoli soavtorstva sebi, če ni pomembneje prispeval k raziskavi, o kateri govori članek. Noben znanstvenik ali znanstvenica naj ne privoli v prisilo predstavnikov akademske ustanove, podjetja, vladne agencije ali katere koli druge osebe, da bi si prisvojili soavtorstvo za raziskavo, kjer nimajo pomembnih zaslug; prav tako naj noben znanstvenik ne dovoli uporabe neupravičenega soavtorstva v zameno za kakršno koli darilo ali drugo podkupnino. Nihče naj kakor koli ne sili znanstvenika, da bi bil kot soavtor pripisan kdor koli, ki ni pomembno prispeval k raziskavi v članku.

10 člen: Neodvisnost pripadnosti

Danes je veliko znanstvenikov zaposlenih na podlagi kratkoročnih pogodb. S prekinitvijo pogodbe o zaposlitvi ugasne tudi akademska pripadnost. Med uredniškimi odbori pogosto prevladuje politika, da se člankov tistih brez akademske ali komercialne pripadnosti ne objavlja. Zaradi takšne izključenosti znanstvenik nima dostopa do mnogih virov, zmanjšajo se mu tudi možnosti za predstavitev govorov in razprav na konferencah. To nečedno prakso je treba ustaviti. Znanost se ne prepoznava po pripadnosti.

Zaradi umanjkanja pripadnosti akademski ustanovi, znanstvenemu inštitutu, vladnemu ali komercialnemu laboratoriju ali kateri koli drugi organizaciji naj noben znanstvenik ne bo prikrajšan za možnost predstavitve svojih člankov na konferencah, kolokvijih in seminarjih, za objavljanje v katerem koli mediju, za dostop do knjižnic ali znanstvenih publikacij, za udeležbo na znanstvenih simpozijih in za izvajanje predavanj.

11 člen: Prost dostop do znanstvenih informacij

Večina specializiranih znanstvenih knjig in veliko znanstvenih revij ustvarja malo ali nič dobička, tako da jih komercialni založniki niso pripravljeni izdajati brez denarnih prispevkov, ki jih nudijo akademske ustanove, vladne agencije, človekoljubni skladi in podobni. V takšnih okoliščinah bi morali komercialni založniki dovoliti prost dostop do elektronskih različic publikacij in si prizadevati za čim nižjo ceno tiskovin.

Vsi znanstveniki naj si prizadevajo, da bi bili njihovi raziskovalni članki brezplačno dostopni za mednarodno znanstveno skupnost; če ne gre drugače, pa vsaj za minimalno

ceno. Vsi znanstveniki naj se lotijo oprijemljivih ukrepov in ponudijo svoje strokovne knjige po najnižji mogoči ceni, saj bodo znanstvene informacije le na tak način na voljo širši mednarodni znanstveni skupnosti.

12 člen: Etična odgovornost znanstvenikov

Zgodovina priča, da so znanstvena odkritja lahko v rabi tako za dobre kot zle namene: za blagor enih in v pogubo drugih. Ker se napredka znanosti in tehnologije ne da ustaviti, je treba zagotoviti razmere za omejitev zlonamerne rabe. Le demokratično izvoljena vlada brez verskih, rasnih in drugih predsodkov lahko obvaruje civilizacijo. Le demokratično izvoljena vlada, sodišča in odbori lahko obvarujejo pravico do svobodnega znanstvenega ustvarjanja. Danes različne nedemokratične države in totalitarni režimi izvajajo dejavne raziskave na področju jedrske fizike, kemije, virologije, genetskega inženiringa in še kje, z namenom, da bi naredili jedrsko, kemijsko in biološko orožje. Noben znanstvenik naj prostovoljno ne sodeluje z nedemokratičnimi državami in totalitarnimi režimi. Vsak znanstvenik, ki je prisiljen sodelovati pri razvoju orožja za takšne države, mora najti način in sredstva za upočasnitev napredovanja raziskovalnih programov in zmanjšati znanstveni učinek, tako da lahko civilizacija in demokracija na koncu prevladata.

Vsi znanstveniki so moralno odgovorni za svoje znanstvene stvaritve in odkritja. Noben znanstvenik naj samovoljno ne sodeluje pri načrtovanju in izdelavi orožja kakršne koli vrste za kakršno koli nedemokratično državo ali totalitarni režim ali dovoli uporabe svojih znanstvenih veščin in znanja za razvoj česar koli takšnega, kar bi lahko na kakršen koli način ogrožalo človeštvo. Znanstvenik naj živi, kakor veli naslednje reklo: ⟩⟩Sleherna nedemokratična vladavina in kršitev človekovih pravic sta zločin!⟨⟨

Posvetilo (Dedication)

Ta prevod je posvečen Manici, prevajalčevi drugi hčeri in prevajalkini nečakinji, ter ekipi dvigalcev uteži Plamen.

This translation is dedicated to Manica, Translators' second daughter and niece, respectively, and to the weightlifting team Plamen.

V Gornji Radgoni, 14. julija 2013

LETTERS TO PROGRESS IN PHYSICS

On Meta-Epistemic Determination of Quality and Reality in Scientific Creation

(An Address to Those Against Real Science, Scientific Creation, Intellectual Freedom, and Epistemic Culture)

Indranu Suhendro

The Zelmanov Cosmophysical Group
http://www.zelmanov.org

This is an open letter entitled as "On Meta-Epistemic Determination of Quality and Reality in Scientific Creation". An address to those against real science, scientific creation, intellectual freedom, and epistemic culture. Inspired by the Declaration of Academic Freedom.

Suffice it to say once and for all that you — and so many others like you — are not epistemically qualified to assess and categorize in any way my person, my work, nor any of my highly dignified and most devoted colleagues (as profoundly silent and understanding as they are), nor our scientific-philosophical group as a whole, both positively and negatively, whether in whole or in part. Such an attempt — particularly such a smug, narrow, shallow, pseudo-intellectual vacuity, which has foamed and mushroomed throughout certain loose forums, online and offline — is essentially epistemically superficial, hollow, arbitrary, and inauthentic, no matter how much pompous sophistication it displays (by this, I simply mean sophisticated solipsism, verbal and mental, stemming from the widespread, persistent epistemic problem of solipsistic syllogism and syllogistic solipsism). It has nothing whatsoever to do with the determination of Quality (quality-in-itself) and Reality (reality-in-itself) in the realmost sense.

The real tragedy of this world, at large (including academia), consists in the lack of epistemic character; of insight and creation (especially scientific creation); of independence and freedom; of objectivity and universality; of honesty and integrity; of solitude and originality; of "qualic" ideation, imagination, intellection, and identity; of a true sense of epistemicity and existentialism; of the ontic-epistemic unity of sight and sense — in other words, of Quality and Reality. These profound characteristics, throughout history, have never been, and will never be, embodied in the collective majority, let alone the very imitators (in contrast to real creators) and their stooges. These belong only to the truly solitary, independent, authentic few among intellectuals capable of not just filibustering and pan-handling raw fragments of knowledge, but also of critically and figuratively substantiating all types of knowledge and understanding. Such an individual is very, very rare.

If you have never heard, nor comprehended, notorious affairs in science such as the Erasmus affair, the Abel affair, the Galois affair, the Bolyai affair, the Wagener affair, the Dewey affair, the Alfven affair, the Sidis affair, the Pirsig affair, and, most recently, the Arp affair, the Wolfram affair, and the Perelman affair (alongside other such affairs in the annals of art and philosophy); whether you deem yourself a scientist or a lay person, you would better not assert anything potentially misleading in this category, especially publicly. As Michael Crichton once lamented, science is not the same, and should never be equal to, "consensus science" — with consensus (often very falsely, abusively masquarading as "democracy" and "objectivity") often being the first and last hiding place (refuge) for scoundrels, mere biased opportunists and affiliates, and pseudo-scientists —; science is simply about one person (or a few), one thinker, one scientist, being correct (in the sense of expanding horizons), no matter how much public opposition and alienation (e.g. Faustian and Kierkegaardian epistemic alienation) he faces, thus contributing not only to the discovery of new facts, but also to the discovery of new ways of thinking and new landscapes of ideation.

That is why in this passage, I shall very militantly emphasize upon the sublime adjective "epistemic" repeatedly (though I generally do not repeat myself): a truly revolutionary science not only contains a new methodology and a new phenomenology, but also a new epistemology and epistemicity, a new ontology and onticity — it introduces new, vaster, more profound "paint", "brush", "canvas", and "dimension", along with a whole new sketch.

Thus, for instance, using the word "fringe" oversimplifyingly and over-homogenizingly when describing a very peculiar scientist or a scientific group, without ever bothering to base it on correct epistemic qualifications, is slanderous, non-scientific, and non-sensical, far removed from real scientific attitude (whether it is perpetrated by academics and politicians first-hand or by lay people). It is a latent trait of characterless pan-academic memesis and mimicry (e.g. as contrasted with the "mnemonist sense" of the Soviet scientist A. Luria) and of pseudo-objectivity, pseudo-science, and pseudo-skepticism (e.g. in the sense of the sociologist of science M. Truzzi).

Besides, basically there are two kinds of "fringes" (referring to both "mere outsiders" and "those who are self-conscious on the boundary") with respect to the majority ("mob consciousness") in any given domain of thought: 1) the utterly wrong "crackpot" one, which is just basic, quickly self-dispersing non-sense without any significance, and 2) the subtle, mercurial "vortical" one — frontier science laden with extreme originality, creativity, synthesis, and daringness –, which DOES have true, profound, substantial epistemic qualification, novelty, merit, and life (i.e. space and direction) in the sublime heart and vein of science, philosophy, and art.

Without this in the very life of the sciences, all good human endeavors, speculations, and ideas are as good as being suffocated, dwarfened, and nullified, and thus organically dead, instead of epistemically, creatively breathing, living, and winging. It is this cross-roads, frontier-type, revolutionary, vortical kind of science that matters the most in the penultimate, genuine progress of science, let alone all of humanity, a merit to be most fairly appreciated in its own universal time, not simply in a temporary "age" dominated by some contemporaneous power-structures and political interests.

To paraphrase Schopenhauer, every genuine — truly epistemically original and weighty — truth, along with its markedly lone proponents (included are the geniuses and mavericks concerned not with merely "adding color and ice to a pre-existing drink and cup", but with opening new frontiers, dimensions, and grounds entirely), is effervescently conscious of three stages pertaining to the reactionary, abusive behavior of the crowd, the majority, whether practically in power or not: first, it is ignored; second, it is ridiculed, rejected, slandered, and violently opposed; third, it is accepted as "self-evident" — and yet this last phase is often only in conjunction with Oppenheimer's (and Kuhn's) warning, "they (the proponents of fortress status-quo) do not get convinced ever, they simply die first".

In this sense, and only in this sense, there is no such a thing as a "single scientific method". Serious paradigms coexist at the frontiers not as mere parallels and alternatives with respect to each other, but already as profound alternating paradigms.

Genius, one with genuine academic freedom, is the very faculty responsible for novelty in individual scientific creation and collective scientific production, including, inevitably at a very fundamental level, new scientific theories, syntheses, and results as well as new ways of managing science altogether. This is because the structure of scientific revolution takes place simultaneously at methodological, phenomenological, axiological-ethical, epistemological, and even ontological levels. One cannot separate individual scientific creation and collective scientific production from the underlying philosophy and sociology of science. This way, self-aware epistemology serves as the very gradient on the slope of knowledge all the way to the mountain peak of scientific progress and revolution.

Suppression, abuse, slander, and any other kind of ill-treatment done by the majority towards anything intellectually new and blossoming by a minority in this category can truly be likened to child abuse: for here we are dealing with the infancy and growth — as well as the very ground, seeds, roots — of future scientific clarity, superstructures, and foundations.

Science evolves, revolves, snarls, twists, and surmounts on tensed — indeed epistemically intense and maudlin — edges and ridges, on suave pavements and narrow lanes, on lone fulcrums and horizons, as well as in broad day-light and in long stringent evenings, in the silent wet limits of the world, in poignant cracks and labyrinths; and the spirit of scientific revolution, let alone dialectics, is embodied this way, through critical, paradoxical, synthetic, epistemic, universal free thinking. Any form of dogmatic suppression and stymie in science in any epoch (i.e. in antiquity, modernism, post-modernism, and "post-post-modernism") is intolerable, a cumbersome instance which usually easily shows itself perfidiously in cases of epistemically hideous over-funding, over-politicization, over-elitism, over-sycophancy, over-patronizing, and over-establishment.

If one is not uniquely, naturally well-versed in these logico-dialectical strands of thinking, one is simply not a real scientist and creator capable of any profound insight and zenith. Such an attitude should also underlie a real, truly enlightened scientific enterprise and editorship: irrespective of the individual views of the editors and reviewers of a scientific guild, one must allow diverse new ideas to flourish and co-exist (as long as they are true new ideas, and not obvious "pieces of crackpottery", in the minimum epistemic sense). This should naturally, winnowingly manifest spontaneous scientific-epistemic certainty and solidity, far removed from the prevalent type of superficial insecurity, fear, and suppression.

While a scientist, I am also an acutely epistemic artist, independent philosophical mind, keen observer-participant, and free thinker, and this indelible quality wholly underlies my scientific path. Insight, originality, creativity, and solitude are the things that matter the most to me — not mere conformity, suitability, respectability, and normalcy. If I display my work of art (e.g. painting, sculpture, and musical score), and if it is indeed my very own authentic creation and self-conscious novel expression of profundity and eccentricity, I need not list any so-called "references": the object — the work — is ALREADY there in its entirety, and it is lone, universal, and transparent as it is, possessing both a verizon and a horizon. True originality shines through effortlessly, especially as regards scientific creation (and not mere "review" or "documentation"). There is no difference in this matter, whether I create scientifically, artistically, or philosophically: when I create something, I create it in a most comprehensive scientific, artistic, and philosophical sense. This ensures

real quality. Reality alone — and the Universe — is the parameter, not fallible and unqualified observers. It goes without saying that my "predecessors" in this drive naturally include Einstein, who did not bother to do the "administrative non-essentials" (listing so-called "references") in his 1905 and subsequent revolutionary papers, and Wittgenstein, who hardly referred to some other work in his 1918 masterpiece *Tractatus Logico-Philosophicus*.

Pueril, arbitrary comments such as the ones you and the many often perpetrate in a popular forum, and in certain other forums, are but mere psychological detours, infinitely away from real objectivity, verging on typical character assassination and individual abuse. Given a Rembrandt painting, or at least a Modigliani one, or indeed the work of any pan-Renaissance artist, one should not speak of the "person" of the artist in such a cowardly, biased, envious way or hastily resort to ill-chatter, but, first and foremost, one should behold and withhold, witness and withstand, his very art, ALREADY laid bare and transparent for all its mystery and mastery. If one still does not know what one is trying to comprehend or appreciate here, one should at least possess silent humility before the horizon and verizon of things: the qualitative distance between substantial ideas and mere opinions is infinite and asymmetric. It is ethically, universally very lame to form mere borrowed opinions, to downplay certain contributions, and to resort to ad hominem attack, as is often the case. Opinions are mere opinions, not real ideas, let alone absolute truths. I repeat: "Doxa" is never the same as "Eidos". One is here speaking of the determination and qualification of Reality and Quality, i.e. of "unicity" and "qualicity".

Again, certain such popular treatments verging on the immoral and the ethically ill are epistemically very trivial, categorically replete with misleading logical error (non-sequitur), ad hominem attack, individual abuse, hyper-semiotics, hypernarration, oxymornonism, pseudo-science, pseudo-skepticism, pseudo-philosophy, pseudo-objectivity, solipsism, and epistemic shallowness.

You know nothing about us first-hand, absolutely nothing. You have only seen shadows and facades, and have only heard petty rumors, slander, and gossip (while we never seek enemies and pettiness in any case). We protect our individuality and wish to advance common scientific freedom and objectivity so universally much, perhaps "too much", that we rarely enlist "who we are", other than simply delivering our objectives. An objective of ours is not mere "inter-subjectivity", but truly epistemically qualified.

As regards "who we are", we are simply peculiar general relativists and cosmologists as well as core theoreticians and experimentalists. Also, we have never enlisted all our helpers/supporters one by one as well as our real "address" at length — only a decoy tertiary one for mere administrative and convenience purposes, not scientific purposes — for it matters not whether we reveal such things or not. What matters is the science. We are a core body of just a few acutely epistemic- progressive science creators throughout the world. That said, our group has more than one headquarters in the world. What essentially matters is the real scope, puissance, renaissance, and dimension of our scientific productivity and guardianship. We, a unique combination of the "very young" and "very old", epistemically and experientally, are serving science, philosophy, artistry, and humanity with all our strength, in necessary absolute freedom.

Indeed, some of us have had core scientific experiences as far back as the two world wars and the cold war along the contours of history, scientific creation, existential alienation, political turbulence, and cultural-scientific administration. A lot of us have synthesized first-hand the landscapes of both core Soviet and American science, East and West, and beyond. We are neither "big" nor "small"; we are infinite and infinitesimal. We know the world within and without, within-the-within and without-the-without. We alone know who we are. We know history and the human tendencies very well. We truly know where we have come from and where we are heading. We are quintessentially scientific and humanistic.

We do not populate typical non-scientific forums (especially countless on the internet), where mere bipolar, biased opinions are inevitably found in abundance: we are scientists in the most extreme sense of epistemic integrity and predisposition. We do not have time for trinkets, no matter how popular or trendy. We cherish creative solitude, universality, objectivity, independence, and democracy, so uniquely, so intensely, in a single, most variegated meta-epistemic framework, in order to be able to fully, impartially contribute to the betterment of our world in the way we know the most.

Do not bother to respond to this letter: you and so many others are not qualified to do so properly. Doing so shall only reveal, again and again, the very epistemic limitations you have at your core, and hence the very lack of substance lingering therein. Besides, this address is not a mere intellectual rambling or raving, it is simply meant to be a celestial sonnet akin to an ocean symphony and a contrapuntal melody, with "all the secret knowledge of harmony and counterpoint". Now, we shall withdraw into infinite silence, as usual, ever-pugnaciously dwelling in the realm of pure scientific creation.

Thus I hereby declare, once again, all-time individual and collective academic freedom in science, from science, to science, for science.

* * *

Dedicated in the name of truth, beauty, science, creativity, freedom, and genius to Grisha Perelman. And to a much better world rid of the rigid and frigid excess of characterless politics, solipsism, suppression, tyranny, and conformity; a most tranquil, vivid, living world-organism genuinely fond of self-growth and of ideation, individuation, character, liberty, and honesty.

Appendix: Overture on Character and Independence*

Talent warms-up the given (as they say in cookery) and makes it apparent; genius brings something new. But our time lets talent pass for genius. They want to abolish the genius, deify the genius, and let talent forge ahead.

Kierkegaard

Philosophy becomes poetry and science imagination, in the enthusiasm of genius.

Disraeli

In every work of genius, we recognize our own rejected thoughts; they come back to us with a certain alienated majesty.

R. W. Emerson

Genius is the ability to act rightly without precedent — the power to do the right thing the first time.

Elbert Hubbard

Society expresses its sympathy for the geniuses of the past to distract attention from the fact that it has no intention of being sympathetic to the geniuses of the present.

Celia Green

There is in every [such] madman a misunderstood genius whose idea, shining in his head, frightened people, and for whom delirium was the only solution to the strangulation that life had prepared for him.

Antonin Artaud, of Van Gogh

The case with most men is that they go out into life with one or another accidental characteristic of personality of which they say: "Well, this is the way I am. I cannot do otherwise". Then the world gets to work on them and thus the majority of men are ground into conformity. In each generation a small part cling to their "I cannot do otherwise" and lose their minds. Finally there are a very few in each generation who in spite of all life's terrors cling with more and more inwardness to this "I cannot do otherwise". They are the geniuses. Their "I cannot do otherwise" is an infinite thought, for if one were to cling firmly to a finite thought, he would lose his mind.

Kierkegaard

It is easy to live after the world's opinion; it is easy in solitude to live after your own; but the great man is he who, in the midst of the crowd, keeps with perfect sweetness the independence of solitude.

R. W. Emerson

I call that mind free which protects itself against the usurpations of society, which does not cower to human opinion, which feels itself accountable to a higher tribunal than man's, which respects itself too much to be the slave of the many or the few.

Channing

The genius differs from us men in being able to endure isolation, his rank as a genius is proportionate to his strength for enduring isolation, whereas we men are constantly in need of "the others", the herd; we die, or despair, if we are not reassured by being in the herd, of the same opinion as the herd.

Kierkegaard

Talent is hereditary; it may be the common possession of a whole family (e.g. the Bach family); genius is not transmitted; it is never diffused, but is strictly individual.

Otto Weininger

The age does not create the genius it requires. The genius is not the product of his age, is not to be explained by it, and we do him no honour if we attempt to account for him by it ... And as the causes of its appearance do not lie in any one age, so also the consequences are not limited by time. The achievements of genius live for ever, and time cannot change them. By his works a man of genius is granted immortality on the earth, and thus in a threefold manner he has transcended time. His universal comprehension and memory forbid the annihilation of his experiences with the passing of the moment in which each occurred; his birth is independent of his age, and his work never dies.

Otto Weininger

It is the genius in reality and not the other who is the creator of history, for it is only the genius who is outside and unconditioned by history. The great man has a history, the emperor is only a part of history. The great man transcends time; time creates and time destroys the emperor.

Otto Weininger

Genius is the ability to escape the human condition; Humanity is the need to escape.

Q. Uim

Some superior minds are unrecognized because there is no standard by which to weigh them.

Joseph Joubert

Thousands of geniuses live and die undiscovered — either by themselves or by others.

Mark Twain

Geniuses are like thunderstorms. They go against the wind, terrify people, cleanse the air.

Kierkegaard

A genius is one who can do anything except make a living.

Joey Adams

Could we teach taste or genius by rules, they would be no longer taste and genius.

Joshua Reynolds

*Courtesy: Kevin Solway's extensive philosophical library.

Genius is the highest morality, and, therefore, it is every one's duty. Genius is to be attained by a supreme act of the will, in which the whole universe is affirmed in the individual. Genius is something which "men of genius" take upon themselves; it is the greatest exertion and the greatest pride, the greatest misery and the greatest ecstasy to a man. A man may become a genius if he wishes to. But at once it will certainly be said: "Very many men would like very much to be *original geniuses*", and their wish has no effect. But if these men who "would like very much" had a livelier sense of what is signified by their wish, if they were aware that genius is identical with universal responsibility — and until that is grasped it will only be a wish and not a determination — it is highly probable that a very large number of these men would cease to wish to become geniuses.

Otto Weininger

Universality is the distinguishing mark of genius. There is no such thing as a special genius, a genius for mathematics, or for music, or even for chess, but only a universal genius. The genius is a man who knows everything without having learned it.

Otto Weininger

Genius is the capacity for productive reaction against one's training.

Bernard Berenson

It is frequently the tragedy of the great artist, as it is of the great scientist, that he frightens the ordinary man. If he is more than a popular story-teller it may take humanity a generation to absorb and grow accustomed to the new geography with which the scientist or artist presents us. Even then, perhaps only the more imaginative and literate may accept him. Subconsciously the genius is feared as an image breaker; frequently he does not accept the opinions of the mass, or man's opinion of himself.

Loren Eiseley, in "The Mind as Nature"

I swear to you, sirs, that excessive consciousness is a disease — a genuine, absolute disease. For everyday human existence it would more than suffice to have the ordinary share of human consciousness; that is to say, one half, one quarter that that which falls to the lot of a cultivated man in our wretched nineteenth century [...] It would, for instance, be quite enough to have the amount of consciousness by which all the so-called simple, direct people and men of action live.

Fyodor Dostoevsky

Great geniuses have the shortest biographies. Their cousins can tell you nothing about them.

R. W. Emerson

The genius is not a critic of language, but its creator, as he is the creator of all the mental achievements which are the material of culture and which make up the objective mind, the spirit of the peoples. The "timeless" men are those who make history, for history can be made only by those who are not floating with the stream. It is only those who are unconditioned by time who have real value, and whose productions have an enduring force. And the events that become forces of culture become so only because they have an enduring value.

Otto Weininger

Talent, lying in the understanding, is often inherited; genius, being the action of reason or imagination, rarely or never.

Samuel T. Coleridge

When a true genius appears in this world, you may know him by this sign, that the dunces are all in confederacy against him.

Jonathan Swift

Precisely because the tyranny of opinion is such as to make eccentricity a reproach, it is desirable, in order to break through that tyranny, that people should be eccentric. Eccentricity has always abounded when and where strength of character has abounded; and the amount of eccentricity in a society has generally been proportional to the amount of genius, mental vigor, and moral courage it contained. That so few dare to be eccentric marks the chief danger of the time.

John Stuart Mill

Genius is its own reward; for the best that one is, one must necessarily be for oneself... Further, genius consists in the working of the free intellect., and as a consequence the productions of genius serve no useful purpose. The work of genius may be music, philosophy, painting, or poetry; it is nothing for use or profit. To be useless and unprofitable is one of the characteristics of genius; it is their patent of nobility.

Schopenhauer

Great passions are for the great of souls. Great events can only be seen by people who are on a level with them. We think we can have our visions for nothing. We cannot. Even the finest and most self-sacrificing visions have to be paid for. Strangely enough, that is what makes them fine.

Oscar Wilde

Fortunately for us, there have been traitors and there have been heretics, blasphemers, thinkers, investigators, lovers of liberty, men of genius who have given their lives to better the condition of their fellow-men. It may be well enough here to ask the question: What is greatness? A great man adds to the sum of knowledge, extends the horizon of thought, releases souls from the Bastille of fear, crosses unknown and mysterious seas, gives new islands and new continents to the domain of thought, new constellations to the firmament of mind. A great man does not seek applause or place; he seeks for truth; he seeks the road to happiness, and what he ascertains he gives to others. A great man throws pearls before swine, and the swine are sometimes changed to men. If the great had always kept their pearls, vast multitudes would be barbarians now. A great man is a torch in the darkness, a beacon: in

superstition's night, an inspiration and a prophecy. Greatness is not the gift of majorities; it cannot be thrust upon any man; men cannot give it to another; they can give place and power, but not greatness. The place does not make the man, nor the scepter the king. Greatness is from within.

Robert Ingersoll

No one suffers so much as he [the genius] with the people, and, therefore, for the people, with whom he lives. For, in a certain sense, it is certainly only "by suffering" that a man knows. If compassion is not itself clear, abstractly conceivable or visibly symbolic knowledge, it is, at any rate, the strongest impulse for the acquisition of knowledge. It is only by suffering that the genius understands men. And the genius suffers most because he suffers with and in each and all; but he suffers most through his understanding...

Otto Weininger

He is a man of *capacity* who possesses considerable intellectual riches: while he is a man of *genius* who finds out a vein of new ore. Originality is the seeing nature differently from others, and yet as it is in itself. It is not singularity or affectation, but the discovery of new and valuable truth. All the world do not see the whole meaning of any object they have been looking at. Habit blinds them to some things: shortsightedness to others. Every mind is not a gauge and measure of truth. Nature has her surface and her dark recesses. She is deep, obscure, and infinite. It is only minds on whom she makes her fullest impressions that can penetrate her shrine or unveil her Holy of Holies. It is only those whom she has filled with her spirit that have the boldness or the power to reveal her mysteries to others.

William Hazlitt

Genius is present in every age, but the men carrying it within them remain benumbed unless extraordinary events occur to heat up and melt the mass so that it flows forth.

Denis Diderot

The ego of the genius accordingly is simply itself universal comprehension, the center of infinite space; the great man contains the whole universe within himself; genius is the living microcosm. He is not an intricate mosaic, a chemical combination of an infinite number of elements; [...] as to his relation to other men and things must not be taken in that sense; he is everything. In him and through him all psychical manifestations cohere and are real experiences, not an elaborate piece-work, a whole put together from parts in the fashion of science. For the genius the ego is the all, lives as the all; the genius sees nature and all existences as whole; the relations of things flash on him intuitively; he has not to build bridges of stones between them.

Otto Weininger

I made art a philosophy, and philosophy an art: I altered the minds of men and the colour of things: there was nothing I said or did that did not make people wonder... I treated Art as the supreme reality, and life as a mere mode of fiction: I awoke the imagination of my century so that it created myth and legend around me: I summed up all systems in a phrase, and all existence in an epigram.

Oscar Wilde, in *De Profundis*

Submitted on: May 16, 2013 / Accepted on: July 21, 2013

LETTERS TO PROGRESS IN PHYSICS

Simple Explanation for why Parallel-Propagating Photons do not Gravitationally Attract

Raymond Jensen

Dept. of Mathematics and Science, Northern State University, Aberdeen SD, 57401 USA. E-mail: rwjst4@alumni.nd.edu

In this article it is shown that photons of light, when traveling in parallel, do not attract one another gravitationally. This has been shown previously using general relativity, however here it is only assumed a Newtonian approximation to the gravitational attraction between photons. The explanation for the lack of gravitational attraction is simple: as co-moving objects accelerate in parallel, the flow of time is retarded, as observed by a stationary observer, according to special relativity. Hence so is the tendency for the objects to move toward one another. As the velocity of the objects approach c, the time required for the objects to approach one another approaches infinity, and so there is no gravitational attraction between objects which move parallel at the speed of light.

1 Introduction

In 1931 Tolman, Ehrenfest and Podolsky [1] were first to publish studies on how light interacts with light gravitationally. Among other things, they found that when photons move in parallel beams, there is no gravitational attraction between them. The authors did not give a physical explanation for this peculiarity. In 1999, Faraoni and Dumse [2] studied the problem of gravitational attraction between photons and concluded that for photons moving in parallel, the reason for the lack of gravitational attraction is due to an exact cancellation of the gravitomagnetic and gravitoelectric forces between them. Both sets of authors used a linear approximation to the metric to come to their conclusions. Here, we come to the same conclusion, but it is argued that the lack of gravitation can be entirely explained in Minkowski spacetime with assumption of the Newtonian approximation for gravity. This is reasonable, since the gravitational fields between photons can be expected to be very weak.

2 No attraction between parallel photons

Consider two free particles separated by distance x initially at rest in empty space with respect to an observer. The observer will find that after a time interval t, the objects will come together due to their mutual gravitational attraction. Since the objects are regarded to be small, it is sufficient to assume Newtonian mechanics in the calculation of t, however calculation of the exact value is not necessary for the purpose of the argument here.

Next, consider what happens when the two objects are returned to a distance x apart from one another, accelerated to some terminal velocity v perpendicular to x, and then released. Upon release, the objects initially move parallel to one another, with distance x between, but as before, begin to attract, and eventually come together. However, in this instance, the time required for the two objects to come together, in accordance with special relativity, is $t' = t/\sqrt{1 - v^2/c^2} > t$. Thus, according to a stationary observer, it takes longer for the two objects to approach one another, when their center-of-mass frame is moving at some non-zero velocity. Since the factor $1/\sqrt{1 - v^2/c^2} \to \infty$ as $v \to c$, the time required for the two particles to come together as $v \to c$, approaches infinity. The time required for the objects to deviate from their parallel trajectories is hence also infinite. The conclusion here is that for two particles moving at the speed of light, since time propagation in their center-of-mass frame is nonexistent, their gravitational attraction is also nonexistent. Although x was taken to be perpendicular to the direction of propagation, this condition can be relaxed without changing the conclusion of no gravitational attraction.

3 Attraction between coplanar non-parallel photons

In both of the references, the authors found that for non-parallel propagation, the gravitational attraction between photons is non-zero. This can be reasoned, for some simple cases, as follows: suppose the two particles, in this case photons, are returned to their original positions, but upon release, propagate away from one another at a relative angle $2\theta > 0$, according to a stationary observer. Then, the center of mass frame propagates at a velocity $v = c\cos\theta < c$ and so gravitational attraction between photons is retarded by a factor of $1/\sin\theta$, according to a stationary observer. For example, at $2\theta = 180°$, the photons trajectories are antiparallel to one another, and there is no retardation since the center of mass frame is stationary. The same applies for photons converging at these nonzero angles.

Submitted on: September 10, 2013 / Accepted on September 15, 2013

References

1. Tolman R. C., Ehrenfest P., Podolsky B. On the gravitational field produced by light. *Physical Review*, 1931, v. 37, 602–615.
2. Faraoni V., Dumse R. M. The gravitational interaction of light: from weak to strong fields. *General Relativity and Gravitation*, 1999, v. 31 (1), 91–105.

Progress in Physics is an American scientific journal on advanced studies in physics, registered with the Library of Congress (DC, USA): ISSN 1555-5534 (print version) and ISSN 1555-5615 (online version). The journal is peer reviewed and listed in the abstracting and indexing coverage of: Mathematical Reviews of the AMS (USA), DOAJ of Lund University (Sweden), Zentralblatt MATH (Germany), Scientific Commons of the University of St.Gallen (Switzerland), Open-J-Gate (India), Referential Journal of VINITI (Russia), etc. Progress in Physics is an open-access journal published and distributed in accordance with the Budapest Open Initiative: this means that the electronic copies of both full-size version of the journal and the individual papers published therein will always be acessed for reading, download, and copying for any user free of charge. The journal is issued quarterly (four volumes per year).

Electronic version of this journal: http://www.ptep-online.com

Postal address:
Department of Mathematics and Science,
University of New Mexico, 705 Gurley Avenue, Gallup, NM 87301, USA

Printed in the United States of America